Chemistry of peptide fragmentation

Molecular structures by ion spectroscopy

Josipa Grzetic

Printed by IPSKAMP Drukkers
ISBN/EAN 978-94-6259-392-3

Chemistry of peptide fragmentation

Molecular structures by ion spectroscopy

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. Th.L.M. Engelen,
volgens besluit van het college van decanen
in het openbaar te verdedigen op woensdag 26 november 2014
om 13.00 uur precies

door

Josipa Grzetic

geboren op 20 januari 1980
te Belgrado, Joegoslavië

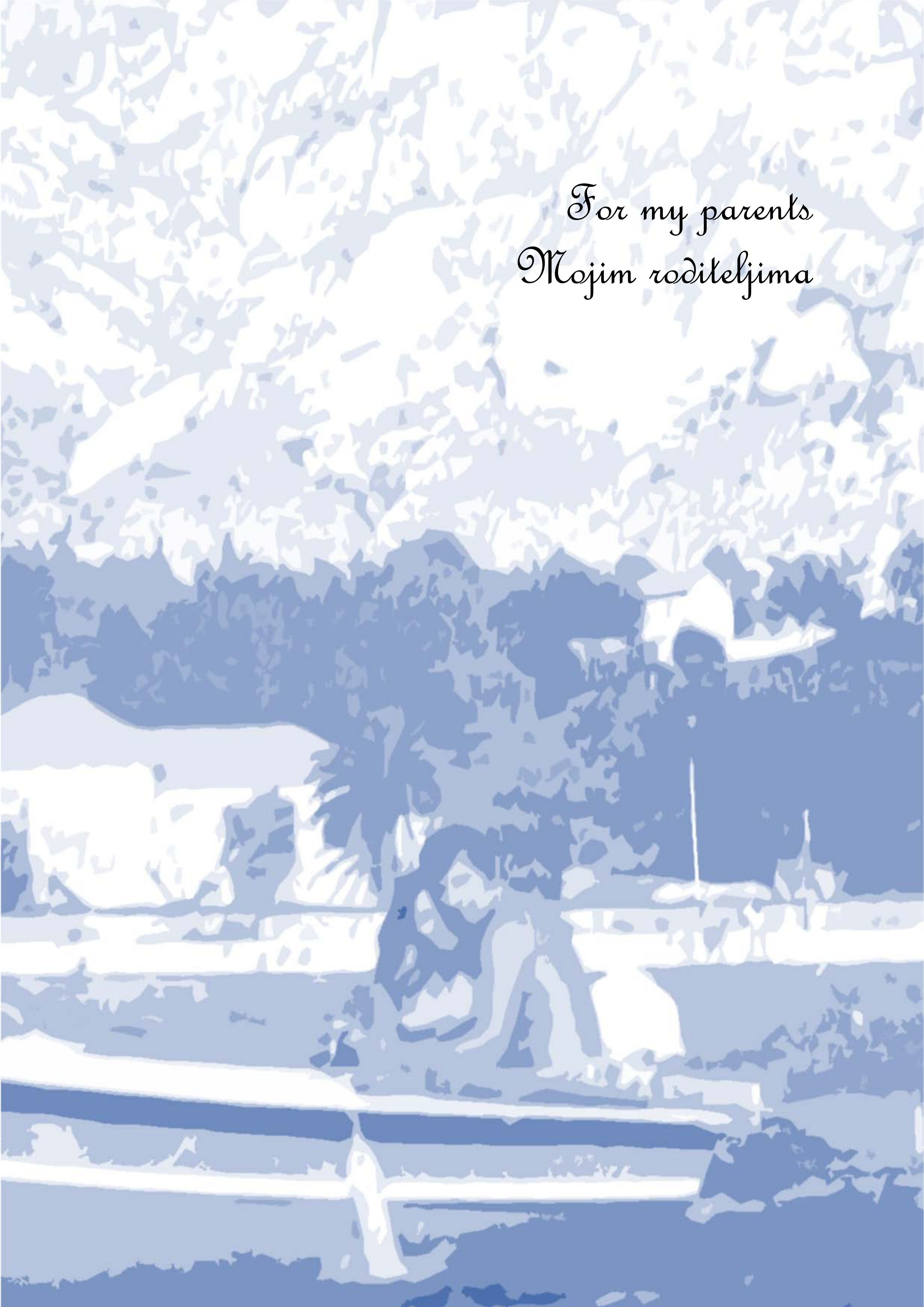
Promotoren

Prof. dr. Jos Oomens
Prof. dr. Wybren J. Buma (Universiteit van Amsterdam)

Manuscriptcommissie

Prof. dr. David H. Parker (voorzitter)
Dr. Britta Redlich
Prof. Dr. Nicolas C. Polfer (University of Florida, Verenigde Staten)

Dit werk maakt deel uit van het onderzoekprogramma van de Stichting voor Fundamenteel Onderzoek der Materie (FOM), die deel uitmaakt van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)



For my parents
Mojim roditeljima

Contents

List of abbreviations

Amino acids	**Ala**	A	Alanine
	Arg	R	Arginine
	Asn	N	Asparagine
	Asp	D	Aspartic acid
	Cys	C	Cysteine
	Glu	E	Glutamic acid
	Gln	Q	Glutamine
	Gly	G	Glycine
	His	H	Histidine
	Ile	I	Isoleucine
	Leu	L	Leucine
	Lys	K	Lysine
	Met	M	Methionine
	Phe	F	Phenylalanine
	Pro	P	Proline
	Ser	S	Serine
	Thr	T	Threonine
	Trp	W	Tryptophan
	Tyr	Y	Tyrosine
	Val	V	Valine

AWE	arbitrary waveform editor
AWG	arbitrary waveform generator
CID	collision induced dissociation
com	centre of mass
DAC	digital to analog converter
DC	direct current
DNA	deoxyribonucleic acid
ECD	electron capture dissociation
ESI	electrospray ionization
FAB	fast atom bombardment
FEL	free electron laser
FID	free induction decay
FT-ICR	Fourier transform ion cyclotron resonance
FT-NMR	Fourier transform nuclear magnetic resonance
GUI	graphical user interface
HPLC	high pressure liquid chromatography
ICR	ion cyclotron resonance
IRMPD	infrared multiple photon dissociation
IVR	intramolecular vibrational (energy) redistribution
LINAC	linear accelerator
m/z	mass to charge ratio
MALDI	matrix assisted laser desorption ionization
MS	mass spectrometry
OPA	optical parametric amplification
OPO	optical parametric oscillation
RF	radio frequency
SASE	self-amplified stimulated emission
SORI	sustained off-resonance irradiation
SWIFT	stored waveform inverse Fourier transformation
TTL	transistor-transistor logic
UV	ultraviolet
VIS	visible

Chapter 1.

Introduction

What are proteins and peptides? Why is it important to analyze their structures? And how is this accomplished? What are the advantages and shortcomings associated with these methods used for structural determination? And where do proteomics and peptide fragmentation chemistry find their place in these questions? And finally, can these techniques be improved? These are some of the questions that will be addressed in the following introductory chapter as they are the key motivation for the work presented in this thesis.

1.1. Proteins and peptides

Proteins are the most plenteous biomolecules occurring in all living cells and in all parts of cells (they make up half the dry weight of an Escherichia coli cell). Proteins are also molecules of high diversity; hundreds of thousands to millions (precise number still unknown) of them, varying in function and ranging in size from small peptides to huge polypeptide complexes, can exist in one living organism. They are directly involved in a vast array of biochemical processes and this diverse functionality is a direct consequence of their conformation (three dimensional structure). Moreover, structurally abnormal proteins (in non-native or misfolded conformation) can lose their native function or even gain toxic effects, further disrupting the normal function of cells, tissues and organs. In medicine this protein-based phenomenon is called *proteopathy* and includes diseases such as Alzheimer's disease, Parkinson's disease, prion disease, amyloidosis and many others. Hence, knowledge of protein structure is of pivotal importance and therefore the subject of numerous investigations.

Remarkable is the fact that all proteins in a living cell, with strikingly different properties, are generally composed of only twenty amino acids. Enzymes (highly selective catalysts, greatly accelerating both the rate and specificity of metabolic reactions), hormones (a class of metabolic regulatory biochemicals), antibodies (part of the immune system), transporters (of nutrients and other molecules), as well as feathers, spider web, hair, horns, antibiotics, mushroom poison and snake venom are examples of proteins, fulfilling numerous biochemical roles.

Twenty simple monomer blocks, encoded by the universal genetic code, provide the key to the architecture of the thousands of distinct proteins. All proteins, whether from the simplest bacteria or the most complex forms of life, are chains formed by combination of amino acids covalently linked in a characteristic linear sequence (Figure 1.1).

Amino acids consist of a central α-carbon to which one hydrogen atom, an amine group and a carboxylate moiety are bonded. The twenty essential amino acids differ by the fourth α-carbon substituent, which is the amino acid side chain (Figure 1.2). Four different substituents make the α-carbon chiral and it can exist in either the L or D configuration (chirality is a structural feature of a molecule that makes it impossible to superimpose it onto its mirror image; one’s left and right hands are examples of chiral objects (Figure 1.2, a). The vast majority of amino acid residues in natural peptides and proteins are L configured. D-Amino acids were found as constituents of only a few bacterial proteins, as well as in a few natural peptide antibiotics. The reason these molecules have such uniform handedness, or "chirality", is not yet known, but there is no shortage of theories on the subject; biological chiral asymmetry is a result of environmental external conditions (light, gravity…), or an incident during evolution that was amplified and propagated over time, or

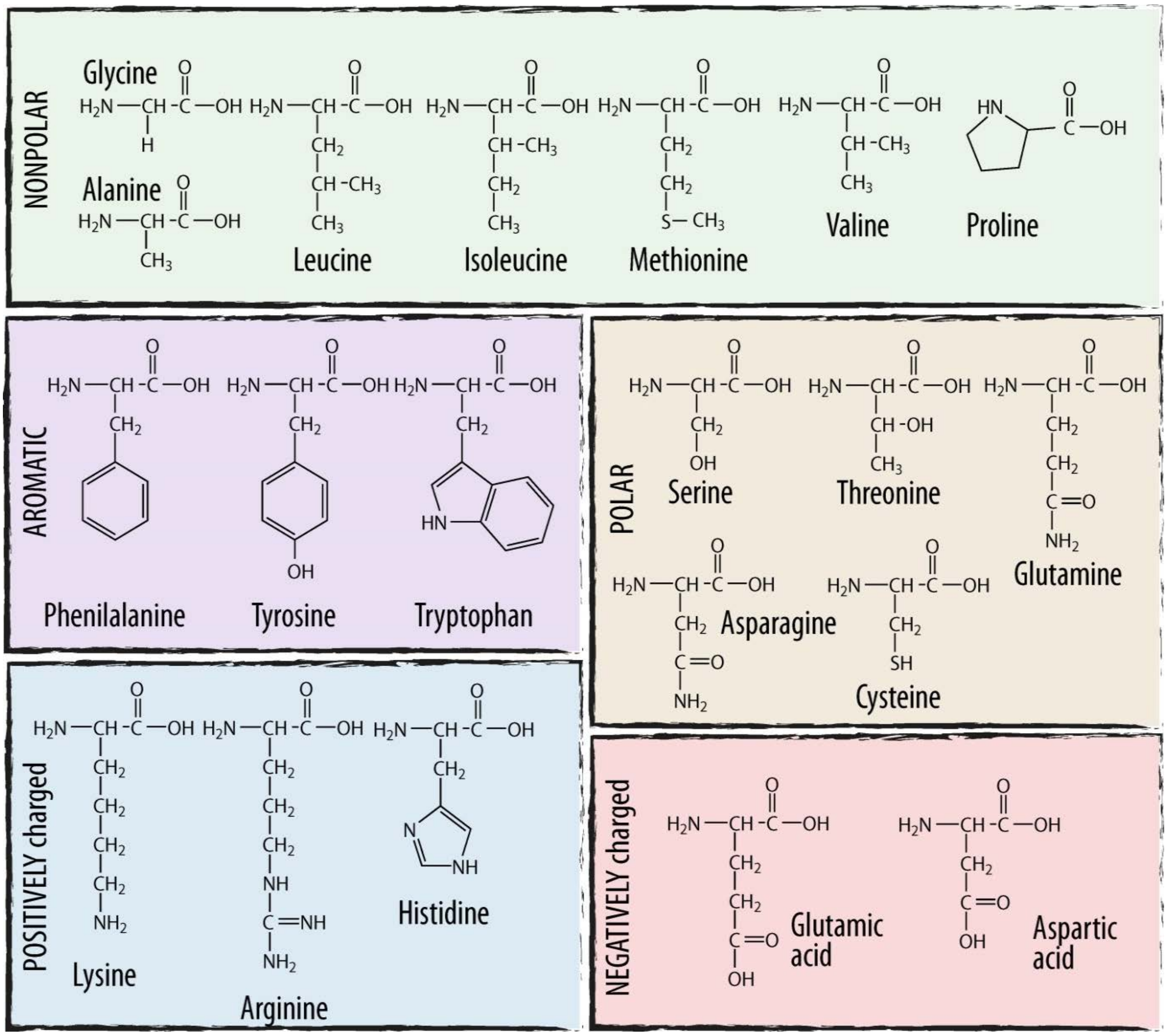


Figure 1.1. **Twenty standard amino acids of proteins.** They can be classified by the chemical properties of the side chain –R, in particular by their polarity or tendency to interact with water at biological pH. This classification is grouping them, as presented, into : nonpolar, aromatic, polar, and positively and negatively charged amino acids.

perhaps L-amino acids were introduced by extraterrestrial bodies (meteorites). The answer to this scientific question would also be a part of an even more popular one "Where does life on earth come from?". Nevertheless, although the complex problem of biomolecules' homochirality is still open, one consequence is well established: this feature of amino acids imparts stereospecificity to biochemical processes.

Two amino acids can undergo a condensation reaction whereby the amino group of one amino acid performs a nucleophilic attack and displaces the hydroxyl group from the carboxylic acid moiety of another amino acid, so that a peptide bond is formed (see Figure 1.3a). Delocalization of the lone pair electrons on the nitrogen atom gives the amide moiety partial double bond character (Figure 1.3b), making the peptide bond (almost) planar. As a result, it can occur in either the cis or trans configuration (Figure 1.4). In most peptide bonds (up to 99.9% of peptide bonds), the trans form is highly preferred energetically. This is partly a result of steric interference of the amino acid side chains, which causes the cis

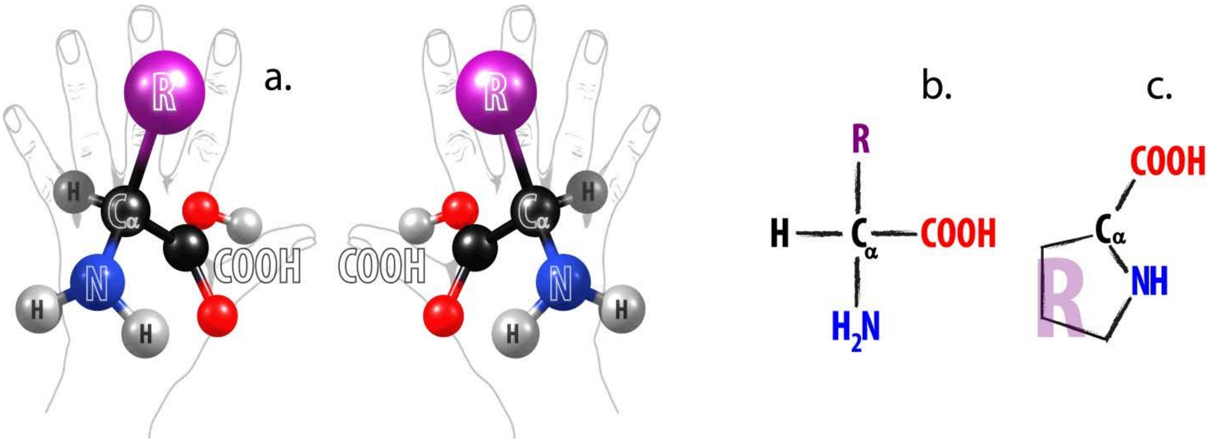


Figure 1.2. (a) Stereoisomers in amino acids - Representation of chirality or handedness (the word chirality is derived from the Greek, -kheir, which means - hand, a familiar chiral object) - the two presented amino acids are enantiomers. **(b)(c) General structural formula of an amino acid**- Nineteen standard amino acids have the same basic structure, they all consist of a carboxylic group (red), amino group (blue), amino acid side chain (R, violet) and the hydrogen atom attached to the Cα (b) the structural formula of proline (c) an exception from the general amino acid structure since it is an amino acid with a unique cyclic structure and secondary amine at the N-terminus.

configuration to be approximately 8 kJ/mol higher in energy than the corresponding trans isomer. Two amino acids linked by a peptide bond are referred to as a dipeptide. If more than a few amino acids are in a chain, the structure is called a polypeptide. Proteins are molecules that consist of one or more polypeptide chains. The terms “protein” and “polypeptide” can sometimes be used interchangeably and there is no clear distinction.

In a peptide, the amino acid with the free amino group is the amino- or N-terminal residue and represents the N-terminus of the peptide. The amino acid at the other peptide end is the carboxyl- or C-terminal residue and is the C-terminus of the peptide. The repeating units of $(\text{-C}_\alpha\text{-peptide bond-})_n$ in a polymer peptide chain is denoted as the peptide backbone (Figure 1.5)

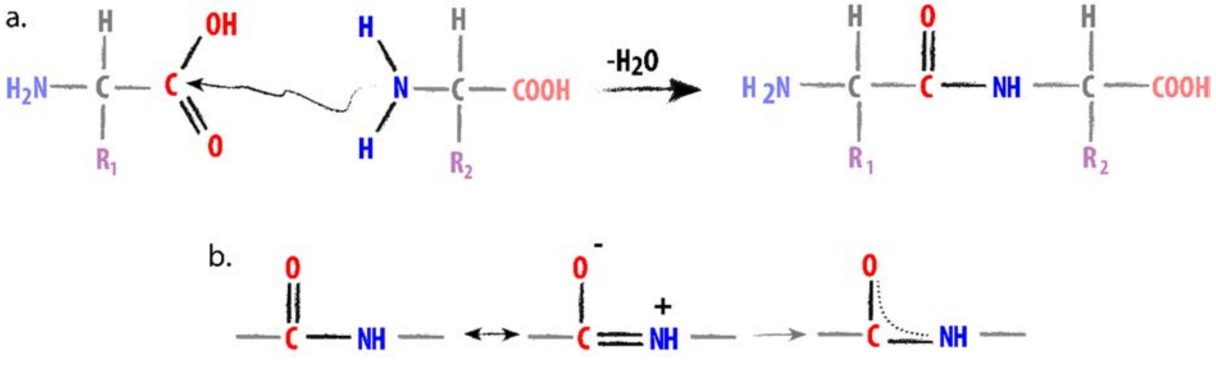


Figure 1.3. **(a) Formation of a peptide bond by condensation**. The amino group of one amino acid acts as a nucleophile and displaces the hydroxyl group of another amino acid forming a peptide bond. **(b) The double bond character of a peptide linkage** is a consequence of resonance delocalization over the amide nitrogen (blue) and carbonyl group (red).

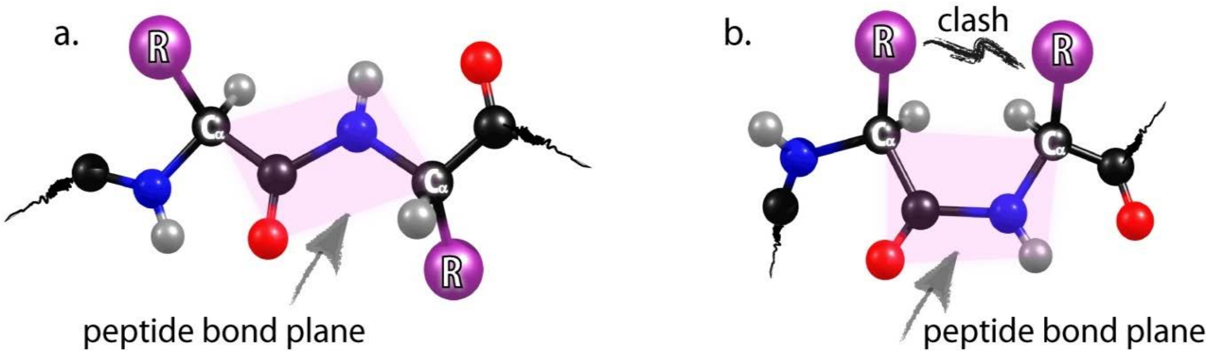
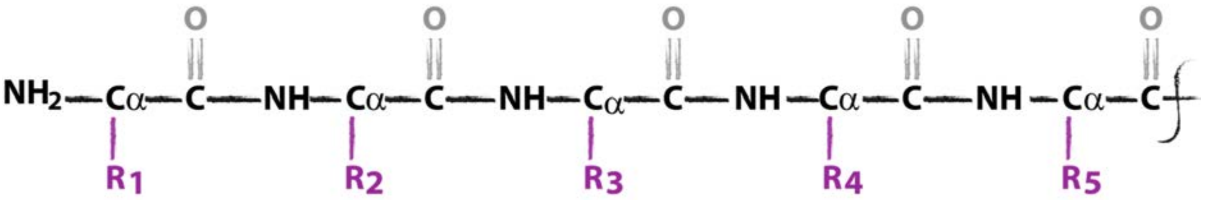


Figure 1.4. (a) Trans peptide bond and (b) cis peptide bond. Due to resonance, a peptide bond has a rigid, planar character. This feature of a peptide bond limits the range of conformations that can be adopted by a polypeptide chain. Peptide bonds, with few exceptions (which will be discussed in later chapters), adopt the trans conformation in which successive C_α atoms are on opposite sides of the peptide bond joining them. The cis configuration induces steric repulsion of side chains (-R) and is therefore energetically unfavorable.



$$NH_2-\underset{R_1}{\underset{|}{C\alpha}}-\overset{O}{\overset{\|}{C}}-NH-\underset{R_2}{\underset{|}{C\alpha}}-\overset{O}{\overset{\|}{C}}-NH-\underset{R_3}{\underset{|}{C\alpha}}-\overset{O}{\overset{\|}{C}}-NH-\underset{R_4}{\underset{|}{C\alpha}}-\overset{O}{\overset{\|}{C}}-NH-\underset{R_5}{\underset{|}{C\alpha}}-\overset{O}{\overset{\|}{C}}-$$

Figure 1.5. Presentation of a polypeptide chain. On the left side is the N-terminus and on the right side is the C-terminus. A peptide backbone that consists of (-C_α-peptide bond-)$_n$ repeating units is presented in black.

1.2. Protein structure

Protein structures built of only amino acid chains are called simple proteins. However, proteins can contain other molecular structures than just amino acids. In such a case, they are called conjugated proteins, and the non-amino acid part of the protein macromolecule is called a prosthetic group. An example is hemoglobin, an iron containing oxygen transport metalloprotein (Figure 1.6). Usually the prosthetic group plays a functional role in the protein. However, the prosthetic group cannot perform the required physiological function alone. Protein structure is commonly described in a four level hierarchy (Figure 1.7): primary structure – describing all covalent (peptide) bonds and representing the amino acid sequence, secondary – describing all stable arrangements and local conformations of the peptide backbone as a result of intramolecular hydrogen bonding and steric interactions (alpha helix or beta sheets), and tertiary – referring to all aspects of the protein folding and incorporation of prosthetic groups. Finally, quaternary protein structure is the arrangement of two or more polypeptide chains together in a bound complex.

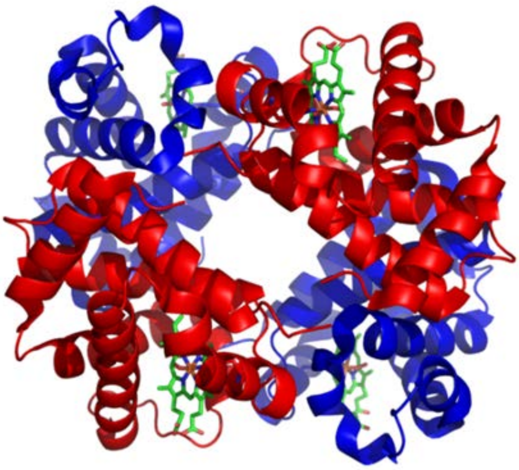


Figure 1.6. Structure of human hemoglobin. Hemoglobin is an oxygen transporting protein in vertebrates. It consists of four polypeptide chains – two identical α-subunits (in red) and two identical β-subunits (in blue). Each of the subunits contains a heme cofactor (in green) each of which contains one iron ion. The iron ion is the site of molecular oxygen binding. Sickle-cell anemia is caused by a point mutation in the β-globin chain of hemoglobin, causing the hydrophilic amino acid glutamic acid to be replaced with the hydrophobic amino acid valine at the sixth position. Sickle-cell anemia is a good (and one of many) example of the protein structure-function correlation.

Proteins differ from one another primarily in their amino acid sequences, which are dictated by the nucleotide sequence of their coding genes. Folding of amino acid arrays into a specific three-dimensional protein structure is what determines its activity. The function of a protein is a direct consequence of its structure. Obviously, the structure of a protein is strongly affected by its amino acid sequence. Determination of the three-dimensional structure of a protein is accomplished using X-ray crystallography or nuclear magnetic resonance (NMR) techniques. There are pros and cons for both of these analytical methods. For example, for a protein to be analyzed using X-ray diffraction, a crystal sample is needed. However, not all proteins, especially hydrophobic membrane proteins, can be crystallized and generally no x-ray structure is available for them. NMR techniques work well, but only for smaller proteins (up to 20 kDa). Both methods require fairly large amounts of purified protein which are not always available. Clearly, there is a constant interest for improvement of known techniques as well as in discovery of new ones.

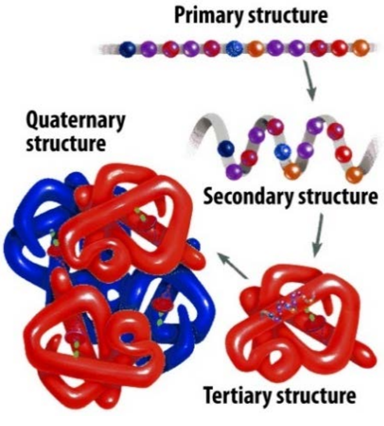


Figure 1.7. Four level hierarchy of protein structure The primary structure is the linear sequence of amino acid residues. Secondary structure is the local spatial arrangement of a polypeptide backbone not taking into account conformation of the side chains. Tertiary structure refers to the three-dimensional structure of the entire polypeptide chain. When one protein consists of more polypeptide chains bound into a multisubunit complex through non-covalent interactions, then three-dimensional structure of that protein is explained with quaternary structure.

1.3. Genomics and Proteomics

Sequencing the human genome (a genome is an organism's complete set of DNA, including all of its genes) initiated a change in how proteins can be studied, since the functional manifestation of DNA is a protein[1]. Development of fast DNA sequencing methods and development of techniques for isolating genes make it possible to deduce the amino acid sequence of a polypeptide by determining the sequence of nucleotides in the gene that codes for it (Figure 1.8). Complete genomes are reported for many organisms, from simple viruses and bacteria to the human genome. It remains puzzling that many genes that have been discovered, encode for proteins without any known function.

On the time scale of all biological processes except evolution, the information of a genomic DNA sequence is static, like a library. Thus, analyzing proteins using only genetic information is insufficient. Importantly, after the information from a gene has been translated into an amino acid sequence, i.e. into a peptide, that peptide may undergo further chemical changes such as phosphorylation, glycosylation, hydroxylation or attachment of prosthetic groups , which are generally referred to as post-translational modifications (PTMs). Other methods must then be applied for a more complete identification of a protein structure. The term proteome ("proteome" is derived from PROTEins expressed by a genOME) has been introduced to describe the complete protein population of a living organism, which is complementary to its entire genome.

Proteomics is the extensive study of a specific proteome, including information on protein abundances, their alterations and modifications, along with their interacting partners and networks. Proteomics has evolved from genomics and the successful sequencing and mapping of the genomes of a wide variety of organisms, including humans. A proteome reflects the cellular state and is much more complex than the genome because each protein can be chemically modified in different ways after biosynthesis/translation. Moreover, the proteome is very dynamic, while the genome remains unchanged over the lifetime of the

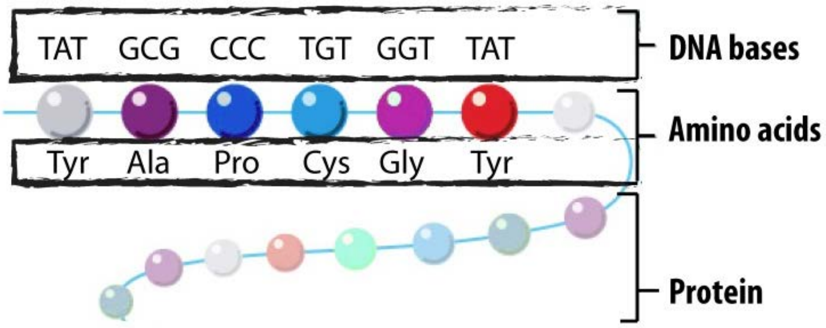


Figure 1.8. Translation of genetic information. The genetic code is the set of rules by which information encoded within genetic material (the DNA or mRNA sequences) is translated into proteins. Three DNA bases form a code called a codon that stands for one amino acid to be incorporated into the final product – the protein. This is obviously just a schematic overview of a very complicated bio-machinery that synthesizes proteins in every living cell.

organism; in contrast, the proteome may vary depending on external conditions, reflecting the life of a cell or a living organism.

From a methodological standpoint, proteomics embodies a collection of various technical disciplines; it includes cell imaging by light or electron microscopy, genetic readout experiments, array or chip experiments and de novo peptide sequencing. **Mass spectrometry** is a major tool in protein identification and structural proteomics.[2-4]

1.4. Mass spectrometry-based proteomics and peptide sequencing

During the 1990's, advances in mass spectrometry, most notably soft ionization methods such as electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI), revolutionized protein chemistry and fundamentally changed the analysis of proteins. These methods made it possible to overcome problems ionizing fragile, large and non-volatile biological molecules such as proteins and peptides[5,6].

Until some twenty years ago, protein sequencing (i.e. identifying the amino acid sequence or primary peptide structure) required one to use a chemical approach known as Edman degradation[7]. Large amounts of purified protein sample were required, as well as hours or days, to sequence a protein of average size. Mass spectrometry (MS) methods replaced Edman degradation because of the increased sensitivity, speed, reliability and lower purification requirements for the sample.

The basic principle of mass spectrometry-based proteomics experiments is presented in Figure 1.9. This technique will be briefly discussed here. First, proteins to be analyzed are isolated from cells/tissue, followed by protein mixture fractionation by gel electrophoresis (protein separation). Isolated proteins are then submitted to proteolytic enzyme digestion, which cuts proteins at the position of specific amino acid residues and results in a mixture of peptides. Peptides from this mixture are then separated by one or more stages of high-pressure liquid chromatography and typically eluted into an electrospray ionization source. A mass spectrum (MS) is recorded for each eluted peptide fraction as well as a fragmentation mass spectrum (MS/MS spectrum) for each of the selected peptides. Peptides are fragmented (as presented in Figure 1.9) by collision induced dissociation (CID), in which the mass selected peptide undergoes energetic collisions with a background gas (usually nitrogen, helium or argon). Other means to induce peptide fragmentation are also available and will be briefly discussed later. Acquiring an MS and an MS/MS spectrum usually requires times on the order of seconds. The objective of MS/MS is to produce a collection of peptide fragment ions that differ in mass by a single amino acid residue, so that the amino acid sequence of the precursor peptide can be reconstructed. The unique

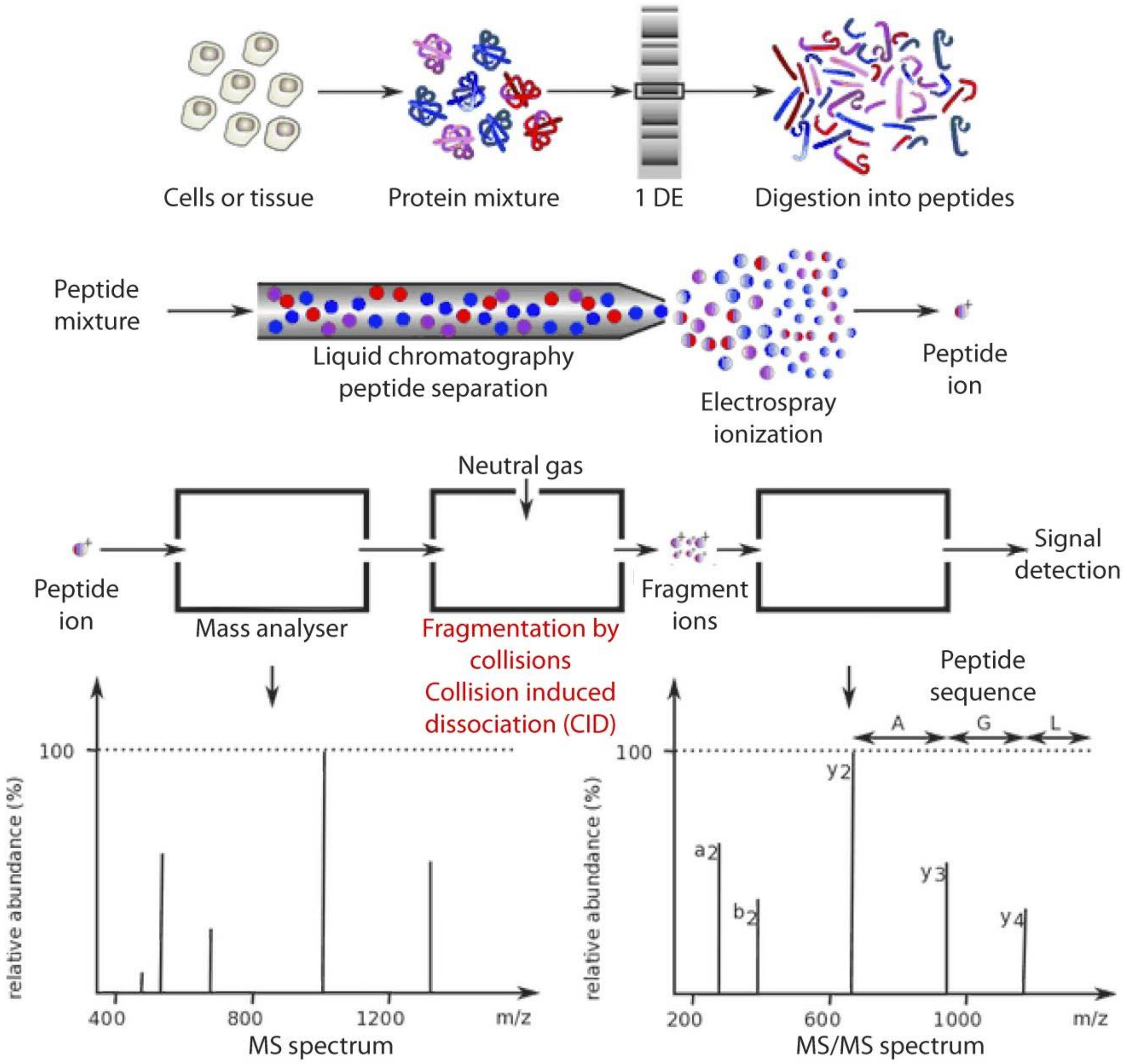


Figure 1.9. Basic principle of MS based proteomics experiment. A protein sample is prepared from a biological source. Proteins are fractionated from the mixture typically by established biochemical methods, such as gel filtration or electrophoresis (that separates macromolecules based on their mass or charge). Protein digestion is accomplished by exposing a (selected) protein to a digestive enzyme such as trypsin. Trypsin cleaves the protein backbone mainly at the C-terminal side of the amino acids lysine and arginine, so that a set of smaller peptides is produced. The generated peptide mixture is then fractionated routinely using high pressure liquid chromatography (HPLC), which is followed by recording the mass spectra of both the precursor ion (MS) and its fragmentation products (MS/MS). Finally, peptide-sequence information is obtained by matching the experimental fragment masses against protein databases using one of various database-searching programs.

molecular weight of each of the 20 naturally occurring amino acids (only leucine and isoleucine are isobaric) makes MS-based proteomics possible. This data is then stored and compared with protein sequence databases. As a final result, peptides and proteins originating from the original sample are identified and possibly quantified.

The two most common strategies in MS based proteomics are top-down and bottom-up proteomics[8]. In top-down proteomics[9], intact proteins are dissociated in a mass spectrometer, i.e. no digestion is done prior to mass analysis. If a sufficient number of informative fragment peptide ions are obtained, this type of analysis can provide accurate information about the original amino acid sequence of a protein and can reveal many of its modifications[9]. However, sequence information may be incomplete because of the largely differing propensities for cleavage at each of the peptide linkages. Moreover, the mass spectra are typically very congested and (expensive) high-resolution instruments (such as FTICR-MS) are required for top-down. The alternative method is the bottom up approach where proteins of interest are first digested with a proteolytic enzyme, commonly trypsin (Figure 1.9). Ideally, those peptide fragments should be assembled into the original amino acid sequence. However, the bottom-up strategy requires higher amounts of sample and is much more time consuming due to the specific demands in sample preparation. Thus, unfortunately, these methods sometimes fail to yield useful sequence information, analogous to having a puzzle with pieces missing - sequence coverage is incomplete.

Results of both bottom-up and top-down protein identification experiments can be used in combination with genomics and/or protein sequence databases to identify the amino acid sequence. Approaches that do not make use of such database comparisons are referred to as de novo sequencing. De novo peptide sequencing is used to identify proteins, often on samples which failed to give a conclusive identification by peptide mass fingerprinting or for samples from organisms for which the genome has not been sequenced. The down side of this technique is that, as de novo sequencing is based only on mass and two amino acids have identical masses (leucine and isoleucine), accurate sequencing can be difficult. Usually, digestion by several different enzymes is needed to obtain a complete sequence coverage, which makes de novo sequencing much more technically demanding and time consuming.

Given the large number of potential proteins in a given system, the interpretation of the peptide fragmentation spectra is handled almost entirely by peptide identification algorithms. Those algorithms use a variety of methods to generate theoretical spectra for a given amino acid sequence to be matched to a recorded MS/MS spectrum. Programs designed to produce sequence information from spectra could be improved if we could more readily predict spectra from a sequence. However, interpretation of the fundamentally complex fragmentation chemistry of peptides is simplified in all algorithms used. For example, it has been shown that many specific peptide cleavages are strongly influenced by side chain chemistry[10,11]. Currently available algorithms either totally ignore or minimally take into account the chemical properties of the different side chains of the 20 amino acid residues. The reason why such residue specific fragmentation information is not incorporated in existing algorithms is mainly the poor fundamental understanding of peptide fragmentation chemistry. In order to more completely exploit the information contained in protein/peptide MS^n experiments, incorporation of a more comprehensive description of peptide fragmentation mechanisms, including both models of non-standard fragmentation pathways and relative fragment-ion intensity information, is required[12,13]. Ongoing studies on the fragmentation of peptides are anticipated to provide a deeper understanding of collected MS data, which could lead to better bioinformatics algorithms for MS-based proteomics.

1.5. Peptide fragmentation chemistry

Early mass spectrometry was useful only for small, volatile compounds. Later, soft ionization techniques such as fast atom bombardment (FAB)[14], electrospray ionization (ESI)[15,16] and matrix assisted laser desorption ionization (MALDI)[6] extended the range of applicability of mass spectrometry to thermally labile compounds. These ionization methods yield precursor ions mainly without any fragmentation. Structural information on the ion can be determined from the fragmentation behavior of the ion. Tandem mass spectrometry involves activation of a precursor ion formed in the ion source followed by mass analysis of the fragments formed. Currently there are several methods of ion activation and they will be discussed here just briefly. CID can be used, where fragmentation is commonly achieved by low-energy collisions with a non-reactive gas[17]; IRMPD – infrared multiphoton dissociation constitutes the irradiation of ions with IR photons leading to dissociation[18]; ECD – electron capture dissociation – exposes the trapped ions to near-thermal electrons and the recombination energy upon capture can dissociate the ion[19] and many others[20]. Several of these activation techniques, when combined, can provide complementary information on ion fragmentation in a mass spectrometer. The bulk of sequencing studies are carried out using CID.

Peptide backbone fragments. Under low-energy CID conditions, peptides mainly fragment along the backbone. The observed fragmentation pattern depends on various factors including the amino acid sequence, the ionization methods applied, charge state, as well as the excitation approach used. Other potentially valuable peptide fragmentation pathways include the loss of small neutral molecules such as water or ammonia. Nomenclature of peptide backbone fragmentation is adopted from Roepstorff[21] et al. as

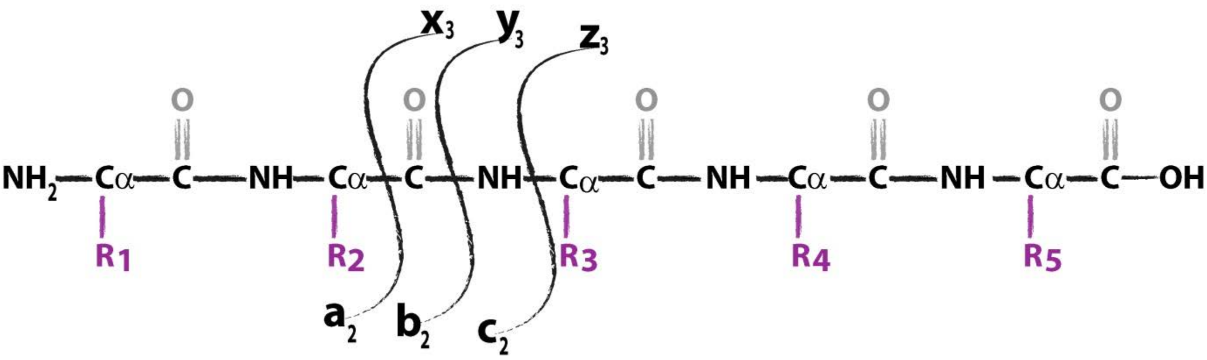


Figure 1.10. Nomenclature of peptide backbone fragments as adopted from Roepstorff. N- terminal fragment ions are classified as either *a*, *b* or *c* while C-terminal ion type is either *x*, *y* or *z*. A subscript indicates the number of residues in the fragment.

illustrated in Figure 1.10. The N-terminal fragment ions are labeled as *a, b,* or *c* depending on which bond of the peptide backbone is actually cleaved. Corresponding C-terminal fragments are labeled as *x, y,* or *z*. Numerical subscripts identify the number of remaining amino acid residues.

An example of a peptide fragmentation spectrum is presented in Figure 1.11. To identify the protein using MS/MS data from protonated peptides, several approaches can be used. First, amino acid residues can be identified from the mass difference of successive fragment ions of the same type (e.g a_n and a_{n-1}). This approach is for example used in *de novo* peptide sequencing. Another approach is the use of peptide sequence tags. A sequence tag is a short string of amino acid mass differences deduced from the fragment spectrum. In this search strategy the peptide fragment spectrum is searched for obvious sequence tags. However, in reality, interpretation of MS data is much more complicated. Some peptides show enhanced or selective fragmentation of peptide bonds at particular amino acid residues and information on the relative fragment peak intensities is completely discarded.

Fragmentation of protonated peptides. Mobile proton theory. The set of rules used to explain MS/MS spectra of peptides were in general derived from experiments. The bond rearrangements involved in the decomposition of an activated ion into two (or more) fragments are triggered by localization of a charge in so called charge-directed fragmentation. Even if almost entirely empirical, the practical success of these “rules” is undeniable[22].

Emerging from a number of studies[23-26], the most comprehensive model used to explain how protonated peptides fragment is the *mobile proton theory*. Having various functional groups, peptides can be protonated at various sites (terminal amino group, amide oxygens and nitrogens, basic side chains) leading to various isomers. An extra proton would preferably settle at the thermodynamically most favorable site. However, as the internal energy of the ions increases upon collisional excitation, energetically less favored protonation sites can become more populated. A proton added to the peptide does not remain fixed at the site of the greatest proton affinity, but it is capable of migrating to positions of lower proton affinity.

Protonation of the amide nitrogen leads to considerable weakening of the peptide bond. This is a result of the lone electron pair of nitrogen being occupied with an extra proton, which disrupts its delocalization over the peptide bond. By contrast, protonation of the amide oxygen would lead to strengthening of the peptide bond.[27,28] Additionally, protonation on the amide nitrogen is not thermodynamically favored compared to protonation on amide oxygens, or on the N-terminal amino group or basic amino acid side chains such as those of arginine and lysine. Thus, for peptide bond cleavage to occur, protonation on the amide nitrogen is likely, while from a thermodynamic perspective, this is not the most favored site. The mobile proton model introduced by the groups of

Wysocki[29] and Gaskell[30] resolves this conflict stating that a proton, upon excitation, can migrate to various protonation sites prior fragmentation.

The mobile proton theory has been verified using deuterium labeling experiments and many other techniques. In these experiments, strong hydrogen/deuterium mixing was observed in a singly deuterated peptide ion upon its activation[31-33]. Mobile proton theory has been successful also in explaining the fragmentation behavior of gaseous proteins; when the number of protons is higher than the number of arginine residues where an extra proton can be sequestered, then a non-selective fragmentation pathway was observed. However, when the number of arginines is larger than the number of extra protons, selective fragmentation, influenced by the aspartic acid amino side chain, is observed[34]. This introduces another possible method of peptide fragmentation.

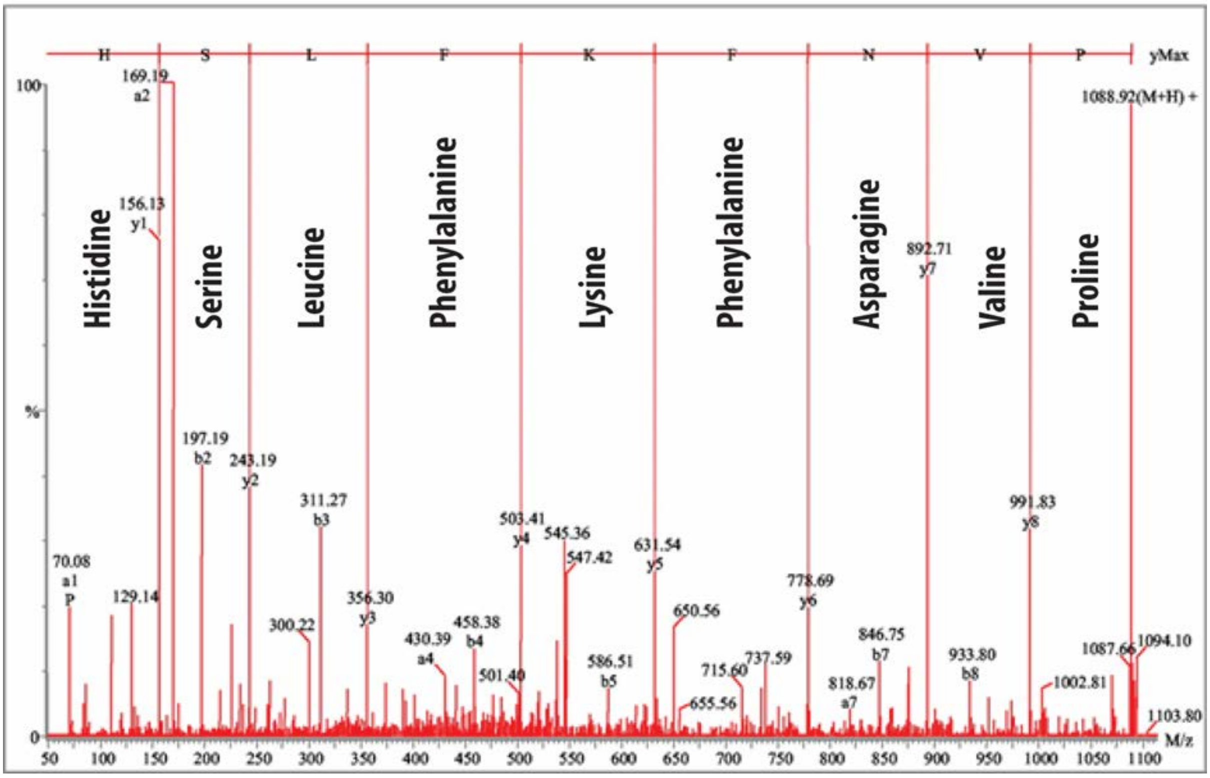


Figure 1.11. Example of an MS/MS spectrum of a protonated peptide of a given amino acid sequence. Difference in ion signal intensities, unassigned peaks and influence of different amino acid side chains are open questions.

In the presence of a highly basic residue such as arginine, more energy is needed to mobilize the proton from the highly-basic arginine side chain to other, less favored protonation sites. Effectively, the proton is sequestered. This type of fragmentation in mass spectrometry is called charge-remote fragmentation and it is more often seen in high-energy collision activated dissociation. Here we can also speculate about the influence of different side chains that can hypothetically perform a nucleophilic attack and induce selective (e.g. amino acid sequence specific) fragmentation of a protonated peptide without proton migration. Such amino acids are, for example, glutamine and asparagine as both have a side-chain amide group. Amide groups feature oxygen and nitrogen atoms and both can act as nucleophiles[35]. This will be discussed in more detail in Chapters 6 and 7.

Pathways in competition model (PIC). The pathways of peptide fragmentation in the competition model are thoroughly explained in a review article by Paizs and Suhai[36]. This PIC model is based on classification of the fragmentation patterns into a simple hierarchy by their mechanistic, energetic, and kinetic descriptions. A kinetic approach is applied to describe the competition of the various pathways. The peptide fragmentation pathways can be grouped according to the scheme presented in Figure 1.12.

Dissociation of protonated peptides can be described as a competition between charge-remote and charge-directed processes, where fragment ions are formed with substantially different probabilities. Protonated peptides in the low fragmentation-energy regime dissociate mainly on charge-directed pathways. The low-energy charge-remote mechanisms correspond to the selective cleavage observed for some of the side chain reactions (aspartic acid, methionine, glutamine or asparagine containing peptides)[37].The charge-directed pathways are further classified as sequence or non-sequence specific dissociation channels. The former produce ions containing information on the primary amino acid sequence of peptides whereas the latter correspond to losses of side chains or small neutrals such as water, ammonia, etc.

Sequence-specific fragments formed from collision induced dissociation of protonated peptides are usually of b-, a- and y-type and result from cleavage of the peptide bond.

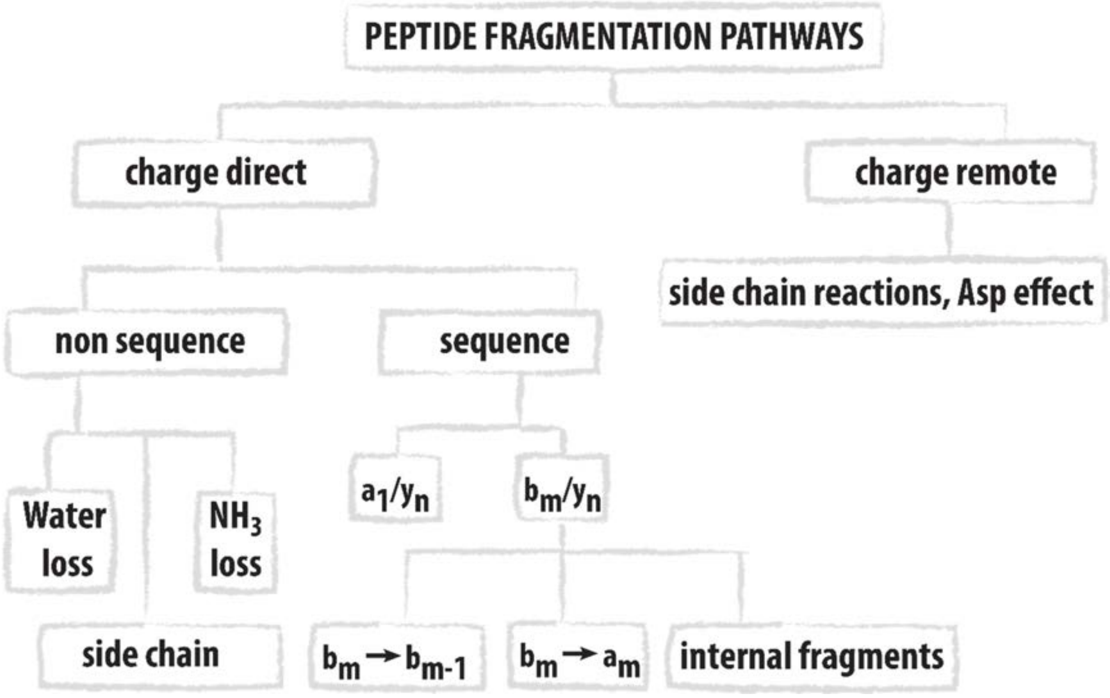


Figure 1.12. Dissociation of protonated peptides described by the pathways in competition (PIC). (Adopted from Ref. 35)

Formation of these fragmentation products involves migration of the additional proton from the energetically most preferred site to one of the amide nitrogens, as explained by mobile proton theory. The carbon atom of the protonated amide moiety becomes a likely target of nucleophilic attack by neighboring electron-rich groups. The b ions formed can fragment further by CO loss to form the corresponding a-ion or to smaller b–ions. Multiple cleavages can also give rise to internal fragments having lost both N- and C-terminal fragments. Competition of these pathways is one of the main factors determining the relative peak intensities in MS/MS spectra of protonated peptides (e.g. which fragments will be observable as ions). The outcome of all the fragmentation reactions above is driven by the thermodynamics and reaction kinetics involved. For instance, the fragment with larger proton affinity (PA) most likely retains the additional proton[38,39].

Possible b/y peptide fragmentation pathways. To better understand the fragmentation reactions of peptides, several groups have examined the mechanisms for the formation of b- and y-type sequence ions. For better comprehension of the dissociation chemistry of peptides, a key issue is "what are the molecular structures of the b- and y-type sequence ions?" While it is believed that y-type ions possess a linear truncated peptide structure, several different isomeric structures have been proposed for b-type sequence ions.[40-45]

According to the mobile proton theory, the added proton migrates from the most basic site in the peptide to one of the amide nitrogen atoms, causing the amide carbonyl carbon to become electron-deficient. Originally, b-ions were assumed to simply possess a linear acylium ion ($-C^+=O$) structure, resulting from cleavage of the amide C-N bond without further rearrangement. Later, Harisson and coworkers suggested that the oxygen atom of the adjacent amide carbonyl acts as a nucleophile during the C-N bond cleavage, leading to the formation of an oxazolone b-type fragment,[39] featuring a five-membered ring structure (Figure 1.13).

However, a peptide often contains several other nucleophilic moieties that could also attack the electrophilic carbon atom, leading to isomeric molecular structures for the b-type fragment. Further, when unprotonated, the N-terminal amine nitrogen atom can act as a strong nucleophile. For oxazolone b-type fragments containing more than 4 or 5 residues, this nitrogen atom has been suggested to attack the carbonyl C-atom of the oxazolone ring, thus forming a macrocyclic peptide. Further dissociation of these cyclic structures can then give rise to the appearance of non-sequence ions in the MS/MS spectrum. The macrocyclic ring may open at a different position along the backbone than where it originally closed, effectively scrambling the original amino acid sequence[46,47]. Experimental evidence for the formation of these cyclic structures is provided by many studies, including spectroscopy[48-50].

The structure of shorter b-type fragments, in particular the b_2 fragment ion, has also been under much debate[44]. A statistical analysis of a large set of tryptic peptide tandem MS spectra revealed that the incidence of this fragment ion is exceptionally high as compared to longer b-fragments, which was suggested to be due to branching reactions partly leading to an altered molecular structure[51]. In particular, the formation of a cyclic dipeptide, i.e. a diketopiperazine structure was suggested. Such structures can form when the N-terminal nitrogen atom acts as the nucleophile attacking the electrophilic amide carbonyl C-atom instead of the adjacent amide carbonyl O-atom (see Figure 1.13). From theoretical studies it is well-known that diketopiperazine structures are in fact lower in energy than the oxazolone motif, although it is also generally assumed to be kinetically disfavored as its formation involves an entropically costly trans-to-cis isomerization of the amide bond.[36,52,53] From spectroscopic studies, the oxazolone structure has been identified as the most common b_2 fragment structure[54-57] while diketopiperazine structures were identified for peptides containing arginine, histidine and from polyprolines.[58-60] Amino acid side chains of relatively high nucleophilicity but relatively poor proton affinity can also play a role in the fragmentation of a peptide bond via a similar nucleophilic attack. Of particular interest are therefore glutamine (Gln) and asparagine (Asn) residues, which feature an amide group in their side chain. The amide nitrogen and oxygen atoms form strong nucleophiles and the proton affinity of the amide group is expected to be lower than that of the N-terminal amine of the peptide.[44] As will be presented in this thesis, instead of oxazolone or diketopiperazine structures, b_2 type fragments from Gln or Asn containing peptides can form cyclic imide structures as a result of side chain involvement in the dissociation reaction (see Figure 1.13).[35]

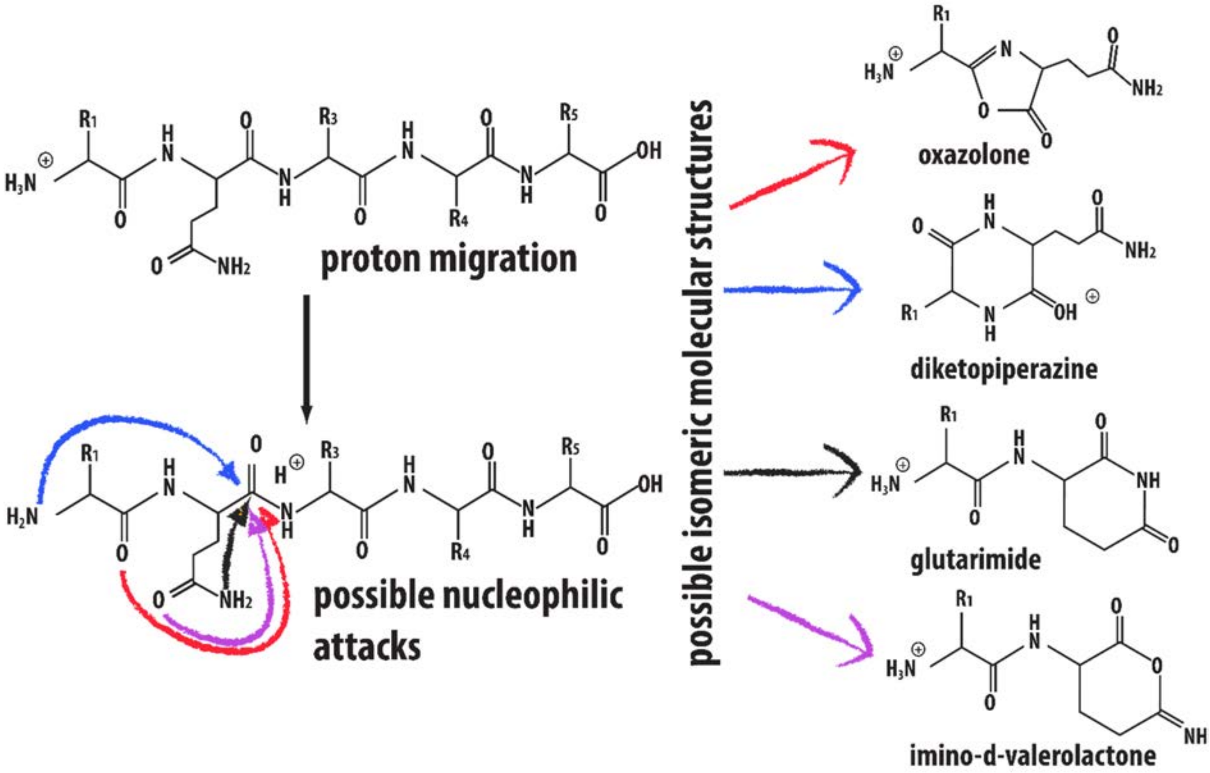


Figure 1.13. Possible dissociation/rearrangement reactions based on the 'mobile proton' model, leading to different b_2-ion structures. In addition to the classic oxazolone and diketopiperazine structures, a Gln residue in the second position allows for the formation of glutarimide (cyclic imide) and imino-δ-valerolactone (cyclic isoimide) structures.

In recent years, IR photo-dissociation spectroscopy of mass-isolated peptide fragments has proven to be an excellent tool to determine the molecular structures of these gas-phase reaction products. Isomeric oxazolone and diketopiperazine b_2 structures are easily distinguished based on their diagnostically different IR spectra. The first b_2-ions investigated with IR spectroscopy contained only Ala and Gly residues and were all found to possess oxazolone structures.[54] Moreover, ion spectroscopy unambiguously identified b_2 ions from three doubly charged tryptic peptides to also exhibit solely oxazolone structures (for possible isomeric structures refer to Figure 1.13).[61,62] The first b_2 ion that was spectroscopically shown to possess a diketopiperazine instead of an oxazolone structure was that of protonated $HisAla_4$.[58] It was suggested that since the added proton in the precursor peptide resides on the His side chain, the N-atom of the N-terminal amino group has greater freedom to carry out the nucleophilic attack leading to the diketopiperazine product ion. This scenario is supported by recent spectroscopic studies on b_2 ions containing an Arg residue, which are also dominantly diketopiperazines.[59]

Fragmentation chemistry of anionic peptides. Although most MS instruments can operate in both positive and negative ion mode, sequencing of peptides is customarily performed on the protonated species. Sequencing a peptide in negative ion mode is feasible and some applications have been reported.[63] Quite evidently, sequence coverage could be increased by analyzing the same peptide in both positive and negative ion mode and using the complementary information.[64] Nonetheless, information on the deprotonated species is barely used in practical proteomics applications, which is perhaps partly due to the fragmentation processes being not as extensively studied.

While backbone fragmentation of protonated peptides generally features a-, b- and y-type ions, negative ion mass spectra of peptides exhibit in addition c and z-type fragment ions. The c-type ions have been observed in CID tandem mass spectra of anionic peptide species containing alkyl [65], acidic,[11,66,67] hydroxyl,[68] basic[69] and aromatic[70] residues, sometimes even with very high abundance. These c-type fragments generated from anionic peptides provide information that is not available from their positive-ion counterparts, and their identification can be used to determine the sequence from the N-terminus and in that way increase the sequence coverage. The IR structural study of these c-type peptide fragment ions is presented in more detail in Chapter 4.

As for protonated peptides, amide bond cleavage is a common peptide fragmentation pattern for deprotonated peptides, leading to the formation of b-type anions. Formation of b-type fragments from deprotonated peptides was first reported by Heerma and co-workers[70-72], and two different fragmentation mechanisms leading to different, but isobaric molecular structures, were proposed, one by Harrison[73] and another by Bowie[74].

What is the structure of c-type peptide anionic fragments? Where is the peptide fragment deprotonated in the gas phase? What are the structures of b-type fragments from deprotonated peptides, do they also form oxazolone or diketopiperazine structures as is well-known for protonated peptides? These and some other questions related to the fragmentation of anionic peptides will be addressed in this thesis by the application of ion spectroscopy.

1.6. Structure determination of biomolecules in the gas phase

Analyzing the structure of biomolecules in the gas phase might seem counterintuitive- biomolecules naturally do not perform their essential function in the gas phase. Rather, they are incorporated in the chemically rich, relatively complicated and highly dynamic environment such as in the living cell. However, in recent decades we have witnessed a progress in the study of biomolecules in the gas phase. There is a three-level motivation for that effort. First, it is of importance to separate intrinsic molecular properties from those properties that result from the interaction with the biological environment. Second, investigation of clusters of biomolecules gives an insight into the fundamental interactions and microsolvation. And third, gas phase data supports, and is used to improve, theoretical quantum chemical methods. In the gas phase, highly charged and highly excited molecules can be studied as well as very cold systems, providing important and complementary information to results from crystallography or solution phase studies.

In order to study biomolecules in the gas phase one has to remove them from their natural environment. Depending on scientific interest, several methods can be applied such as laser desorption and the already mentioned methods MALDI[6] and ESI[15,16]. These modern methods of volatilization have allowed intact biomolecules of virtually any size to be put into the gas phase. Mass spectrometry has become a powerful tool in characterizing biological samples on the molecular level in the gas phase. There are several reasons why MS methods are so advantageous; powerful mass analyzers like Fourier transform mass cyclotron resonance (FTICR)[75] and Orbitrap[76] mass spectrometers allow very accurate mass determination; analyte elemental composition can be determined directly; noncovalent interactions can be preserved in MS experiments giving the opportunity to analyze the stoichiometry of supramolecular assemblies; amino acid sequences can be determined by tandem MS using fragmentation techniques like CID[77] and ECD[78] as described above.

However, MS itself provides very little structural information compared to the combination with other analytical techniques such as spectroscopy. Combination of mass spectrometry and spectroscopy makes possible recording the spectrum of a mass isolated molecular ion. Recorded spectra can then be related to the molecular structure with relatively high confidence when combined with quantum-chemical calculations.

The obstacle in obtaining IR spectra of mass-selected molecular ions is that the density of ions that can be probed in an ion trapping device is limited by Coulomb repulsion of the ions. Therefore, instead of classical direct optical spectroscopy, another type of spectroscopy is used: action spectroscopy. Instead of measuring the ratio of initial and transmitted light intensity, as in direct optical spectroscopy, action spectroscopy measures modifications in the sample as induced by the light absorption. In ion spectroscopy, it is typical to monitor the degree of photo-fragmentation induced by absorption as a function of the wavelength of the radiation.

Emergence of high intensity and tunable free electron lasers (FELs) has made it possible to record the infrared spectra of gas-phase ions[79-81]. The infrared spectrum is obtained as the dependence of the photo-fragmentation yield on the photon frequency. This is the simple definition of infrared multiple-photon dissociation spectroscopy (IRMPD) which is the main experimental tool used in conducting the studies presented in this thesis.

REFERENCES

(1) Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; Smith, H. O.; Yandell, M.; Evans, C. A.; Holt, R. A.; Gocayne, J. D.; Amanatides, P.; Ballew, R. M.; Huson, D. H.; Wortman, J. R.; Zhang, Q.; Kodira, C. D.; Zheng, X. H.; Chen, L.; Skupski, M.; Subramanian, G.; Thomas, P. D.; Zhang, J.; Gabor Miklos, G. L.; Nelson, C.; Broder, S.; Clark, A. G.; Nadeau, J.; McKusick, V. A.; Zinder, N.; Levine, A. J.; Roberts, R. J.; Simon, M.; Slayman, C.; Hunkapiller, M.; Bolanos, R.; Delcher, A.; Dew, I.; Fasulo, D.; Flanigan, M.; Florea, L.; Halpern, A.; Hannenhalli, S.; Kravitz, S.; Levy, S.; Mobarry, C.; Reinert, K.; Remington, K.; Abu-Threideh, J.; Beasley, E.; Biddick, K.; Bonazzi, V.; Brandon, R.; Cargill, M.; Chandramouliswaran, I.; Charlab, R.; Chaturvedi, K.; Deng, Z.; Francesco, V. D.; Dunn, P.; Eilbeck, K.; Evangelista, C.; Gabrielian, A. E.; Gan, W.; Ge, W.; Gong, F.; Gu, Z.; Guan, P.; Heiman, T. J.; Higgins, M. E.; Ji, R.-R.; Ke, Z.; Ketchum, K. A.; Lai, Z.; Lei, Y.; Li, Z.; Li, J.; Liang, Y.; Lin, X.; Lu, F.; Merkulov, G. V.; Milshina, N.; Moore, H. M.; Naik, A. K.; Narayan, V. A.; Neelam, B.; Nusskern, D.; Rusch, D. B.; Salzberg, S.; Shao, W.; Shue, B.; Sun, J.; Wang, Z. Y.; Wang, A.; Wang, X.; Wang, J.; Wei, M.-H.; Wides, R.; Xiao, C.; Yan, C. *The Sequence of the Human Genome* Science **2001**, *291*, 1304.

(2) Nelson, D. L.; Cox, M. M. *Lehninger principles of Biochemistry*; Third ed.; Worth Publisher: New York, **2003.**

(3) Voet, D.; Voet, J. G. *Biochemistry*; Third ed.; Wiley: Danvers, **2004.**

(4) Ruedi, A.; Matthias, M. *Mass spectrometry-based proteomics* Nature **2003**, *422*, 198.

(5) Fenn, J. B.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Electrospray ionization for mass spectrometry of large biomolecules* Science **1989**, *246*, 64.

(6) Karas, M.; Hillenkamp, F. *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons* Anal. Chem. **1988**, *60*, 2299.

(7) Edman, P. *Method for Determination of the Amino Acid Sequence in Peptides* Acta. Chem. Scand. **1950**, *4*, 283.

(8) Chait, B. T. *Mass Spectrometry: Bottom-Up or Top-Down?* Science **2006**, *314*, 65.

(9) Kelleher, N. L. *Peer Reviewed: Top-Down Proteomics* Anal. Chem. **2004**, *76*, 196 A.

(10) Grewal, R. N.; El Aribi, H.; Harrison, A. G.; Siu, K. W. M.; Hopkinson, A. C. *Fragmentation of Protonated Tripeptides: The Proline Effect Revisited* J. Phys. Chem. B **2004**, *108*, 4899.

(11) Harrison, A. G.; Young, A. B. *Fragmentation reactions of deprotonated peptides containing aspartic acid* Int. J. Mass Spectrom. **2006**, *255-256*, 111.

(12) Huang, Y.; Tseng, G. C.; Yuan, S.; Pasa-Tolic, L.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *A Data-Mining Scheme for Identifying Peptide Structural Motifs Responsible for Different MS/MS Fragmentation Intensity Patterns* J. Proteome Res. **2007**, *7*, 70.

(13) Li, W.; Ji, L.; Goya, J.; Tan, G.; Wysocki, V. H. *SQID: An Intensity-Incorporated Protein Identification Algorithm for Tandem Mass Spectrometry* J. Proteome Res. **2011**, *10*, 1593.

(14) Morris, H. R.; Panico, M.; Barber, M.; Bordoli, R. S.; Sedgwick, R. D.; Tyler, A. *Fast atom bombardment: A new mass spectrometric method for peptide sequence analysis* Biochem. Biophys. Res. Commun. **1981**, *101*, 623.

(15) Yamashita, M.; Fenn, J. B. *Electrospray ion source. Another variation on the free-jet theme* J. Phys. Chem. **1984**, *88*, 4451.

(16) Yamashita, M.; Fenn, J. B. *Negative ion production with the electrospray ion source* J. Phys. Chem. **1984**, *88*, 4671.

(17) Mitchell Wells, J.; McLuckey, S. A. In *Methods in enzymology*; Burlingame, A. L., Ed.; Academic Press: 2005; Vol. 402, p 148.

(18) Crowe, M. C.; Brodbelt, J. S. *Infrared multiphoton dissociation (IRMPD) and collisionally activated dissociationof peptides in a quadrupole ion trapwith selective IRMPD of phosphopeptides* J. Am. Soc. Mass Spectrom. **2004**, *15*, 1581.

(19) Coon, J. J. *Collisions or electrons? Protein sequence analysis in the 21st century* Anal. Chem. **2009**, *81*, 3208.

(20) Sleno, L.; Volmer, D. A. *Ion activation methods for tandem mass spectrometry* J. Mass Spectrom. **2004**, *39*, 1091.

(21) Roepstorff P; J., F. *Proposal for a common nomenclature for sequence ions in mass spectra of peptides* Biomed. Mass Spectrom. **1984**, *11*, 601.

(22) McLafferty, F. W.; Turecek, F. *Interpretation of Mass Spectra* Fourth ed.; University Science Books: Mill Valley, CA, **1993**.

(23) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *Mobile and localized protons: A framework for understanding peptide dissociation.* J. Mass Spectrom. **2000**, *35*, 1399.

(24) Harrison, A. G.; Yalcin, T. *Proton mobility in protonated amino acids and peptides* Int. J. Mass Spectrom. Ion Proc. **1997**, *165*, 339.

(25) Jones, J. L.; Dongre, A. R.; Somogyi, A.; Wysocki, V. H. *Sequence Dependence of Peptide Fragmentation Efficiency Curves Determined by Electrospray Ionization/Surface-Induced Dissociation Mass Spectrometry* J. Am. Chem. Soc. **1994**, *116*, 8368.

(26) Boyd, R.; Somogyi, Á. *The Mobile Proton Hypothesis in Fragmentation of Protonated Peptides: A Perspective* J. Am. Soc. Mass Spectrom. **2010**, *21*, 1275.

(27) Somogyi, Á.; Wysocki, V. H.; Mayer, I. *The effect of protonation site on bond strengths in simple peptides: Application of ab initio and modified neglect of differential overlap bond orders and modified neglect of differential overlap energy partitioning* J. Am. Soc. Mass Spectrom. **1994**, *5*, 704.

(28) McCormack, A. L.; Somogyi, A.; Dongre, A. R.; Wysocki, V. H. *Fragmentation of protonated peptides: surface-induced dissociation in conjunction with a quantum mechanical approach* Anal. Chem. **1993**, *65*, 2859.

(29) Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. *Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: Evidence for the mobile proton model* J. Am. Chem. Soc. **1996**, *118*, 8365.

(30) Burlet, O.; Yang, C. Y.; Gaskell, S. J. *Influence of cysteine to cysteic acid oxidation on the collision-activated decomposition of protonated peptides: Evidence for intraionic interactions* J. Am. Soc. Mass Spectrom. **1992**, *3*, 337.

(31) Harrison, A. G.; Yalcin, T. *Proton mobility in protonated amino acids and peptides* Int. J. Mass Spectrom. Ion Proc. **1997**, *165–166*, 339.

(32) Mueller, D. R.; Eckersley, M.; Richter, W. J. *Hydrogen transfer reactions in the formation of "Y+2" sequence ions from protonated peptides* Org. Mass Spectrom. **1988**, *23*, 217.

(33) Johnson, R. S.; Krylov, D.; Walsh, K. A. *Proton mobility within electrosprayed peptide ions* J. Mass Spectrom. **1995**, *30*, 386.

(34) Engel, B. J.; Pan, P.; Reid, G. E.; Wells, J. M.; McLuckey, S. A. *Charge state dependent fragmentation of gaseous protein ions in a quadrupole ion trap: bovine ferri-, ferro-, and apo-cytochrome c* Int. J. Mass Spectrom. Ion Proc. **2002**, *219*, 171.

(35) Grzetic, J.; Oomens, J. *Spectroscopic identification of cyclic imide b_2-ions from peptides containing Gln and Asn residues* J. Am. Soc. Mass Spectrom. **2013**, *24*, 1228.

(36) Paizs, B.; Suhai, S. *Fragmentation pathways of protonated peptides* Mass Spectrom. Rev. **2004**, *24*, 508.

(37) Gu, C. G.; Tsaprailis, G.; Breci, L.; Wysocki, V. H. *Selective gas-phase cleavage at the peptide bond terminal to aspartic acid in fixed-charge derivatives of asp-containing peptides* Anal. Chem. **2000**, *72*, 5804.

(38) Nold, M. J.; Cerda, B. A.; Wesdemiotis, C. *Proton affinities of the N- and C-terminal segments arising upon the dissociation of the amide bond in protonated peptides* J. Am. Soc. Mass Spectrom. **1999**, *10*, 1.

(39) Paizs, B.; Suhai, S. *Towards Understanding the Tandem Mass Spectra of Protonated Oligopeptides. 1: Mechanism of Amide Bond Cleavage* J. Am. Soc. Mass Spectrom. **2004**, *15*, 102.

(40) Harrison, A.; Csizmadia, I.; Tang, T.-H. *Structure and fragmentation of b_2 ions in peptide mass spectra* J. Am. Soc. Mass Spectrom. **2000**, *11*, 427.

(41) Yalcin, T.; Khouw, C.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. *Why Are b ions stable species in peptide spectra?* J. Am. Soc. Mass Spectrom. **1995**, *6*, 1165.

(42) Yalcin, T.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. *The structure and fragmentation of b_n ($n \geq 3$) ions in peptide spectra* J. Am. Soc. Mass Spectrom. **1996**, *7*, 233.

(43) Eckart, K.; Holthausen, M.; Koch, W.; Spiess, J. *Mass spectrometric and quantum mechanical analysis of gas-phase formation, structure, and decomposition of various b_2 ions and their specifically deuterated analogs* J. Am. Soc. Mass. Spectrom. **1998**, *9*, 1002.

(44) Farrugia, J. M.; OHair, R. A. J.; Reid, G. E. *Do all b_2 ions have oxazolone structures? Multistage mass spectrometry and ab initio studies on protonated N-acyl amino acid methyl ester model systems* Int. J. Mass Spectrom. **2001**, *210-211*, 71.

(45) Harrison, A. G. *To b or not to b: the ongoing saga of peptide b ions* Mass Spectrom. Rev. **2009**, *28*, 640.

(46) Harrison, A. G. *Peptide Sequence Scrambling Through Cyclization of b5 Ions* J. Am. Soc. Mass. Spectrom. **2008**, *19*, 1776.

(47) Bleiholder, C.; Osburn, S.; Williams, T. D.; Suhai, S.; Van Stipdonk, M.; Harrison, A. G.; Paizs, B. *Sequence-scrambling fragmentation pathways of protonated peptides* J. Am. Chem. Soc. **2008**, *130*, 17774.

(48) Chen, X. A.; Steill, J. D.; Oomens, J.; Polfer, N. C. *Oxazolone Versus Macrocycle Structures for Leu-Enkephalin* $b_{(2)}$*-*$b_{(4)}$*: Insights from Infrared Multiple-Photon Dissociation Spectroscopy and Gas-Phase Hydrogen/Deuterium Exchange* J. Am. Soc. Mass. Spectrom. **2010**, *21*, 1313.

(49) Chen, X.; Yu, L.; Steill, J. D.; Oomens, J.; Polfer, N. C. *Effect of Peptide Fragment Size on the Propensity of Cyclization in Collision-Induced Dissociation: Oligoglycine* $b_{(2)}$*-*$b_{(8)}$ J. Am. Chem. Soc. **2009**, *131*, 18272.

(50) Tirado, M.; Rutters, J.; Chen, X.; Yeung, A.; van Maarseveen, J.; Eyler, J. R.; Berden, G.; Oomens, J.; Polfer, N. C. *Disfavoring Macrocycle b Fragments by Constraining Torsional Freedom: The "Twisted" Case of QWFGLM* b_6 J. Am. Soc. Mass Spectrom. **2012**, *23*, 475.

(51) Savitski, M. M.; Falth, M.; Fung, Y. M. E.; Adams, C. M.; Zubarev, R. A. *Bifurcating Fragmentation Behavior of Gas-Phase Tryptic Peptide Dications in Collisional Activation* J. Am. Soc. Mass Spectrom. **2008**, *19*, 1755.

(52) Paizs, B.; Suhai, S. *Combined quantum chemical and RRKM modeling of the main fragmentation pathways of protonated GGG. II. Formation of* $b_{(2)}$*,* $y_{(1)}$*, and* $y_{(2)}$ *ions* Rapid Commun. Mass Spectrom. **2002**, *16*, 375.

(53) Armentrout, P. B.; Heaton, A. L. *Thermodynamics and mechanisms of protonated diglycine decomposition: a computational study* J. Am. Soc. Mass Spectrom. **2012**, *23*, 621.

(54) Oomens, J.; Young, S.; Molesworth, S.; van Stipdonk, M. *Spectroscopic evidence for an oxazolone structure of the* b_2 *fragment ion from protonated tri-alanine* J. Am. Soc. Mass Spectrom. **2009**, *20*, 334.

(55) Wang, D.; Gulyuz, K.; Stedwell, C. N.; Polfer, N. C. *Diagnostic NH and OH Vibrations for Oxazolone and Diketopiperazine Structures:* b_2 *from Protonated Triglycine* J. Am. Soc. Mass Spectrom. **2011**, *22*, 1197.

(56) Yoon, S. H.; Chamot-Rooke, J.; Perkins, B. R.; Hilderbrand, A. E.; Poutsma, J. C.; Wysocki, V. H. *IRMPD spectroscopy shows that AGG forms an oxazolone* b_2^+ *ion* J. Am. Chem. Soc. **2008**, *130*, 17644.

(57) Morrison, L.; Somogyi, A.; Wysocki, V. H. *The influence glutamic acid in protonated* $b_3 > b_2$ *formation from VGEIG and related analogs* Int. J. Mass Spectrom. **2012**, *325-327*, 139.

(58) Perkins, B. R.; Chamot-Rooke, J.; Yoon, S. H.; Gucinski, A. C.; Somogyi, A.; Wysocki, V. H. *Evidence of diketopiperazine and oxazolone structures for HA* b_2^+ *ion* J. Am. Chem. Soc. **2009**, *131*, 17528.

(59) Zou, S.; Oomens, J.; Polfer, N. C. *Competition between diketopiperazine and oxazolone formation in water loss products from protonated ArgGly and GlyArg* Int. J. Mass Spectrom. **2012**, *316-318*, 12.

(60) Gucinski, A. C. C.-R., Julia Nicol, Edith Somogyi, ÁrpádWysocki, Vicki H. *Structural Influences on Preferential Oxazolone versus Diketopiperazine* b_2^+ *Ion Formation for Histidine Analogue-Containing Peptides* The Journal of Physical Chemistry A **2012**, *116*, 4296.

(61) Sinha, R. K.; Erlekam, U.; Bythell, B. J.; Paizs, B.; Maitre, P. *Diagnosing the Protonation Site of* b_2 *Peptide Fragment Ions using IRMPD in the X-H (X = O, N, and C) Stretching Region* J. Am. Soc. Mass Spectrom. **2011**, *22*, 1645.

(62) Bythell, B. J.; Erlekam, U.; Paizs, B.; Maitre, P. *Infrared spectroscopy of fragments from doubly protonated tryptic peptides* Chem. Phys. Chem. **2009**, *10*, 883.

(63) *Mass spectrometry and the age of the proteome* J Mass Spectrom **1998**, *33*, 1.

(64) Brinkworth, C. S.; Bowie, J. H.; Bilusich, D.; Tyler, M. J. *The rothein peptides from the skin secretion of Roth's tree frog Litoria rothii. Sequence determination using positive and negative ion electrospray mass spectrometry* Rapid Commun. Mass Spectrom. **2005**, *19*, 2716.

(65) Harrison, A. G. *Sequence-specific fragmentation of deprotonated peptides containing H or alkyl side chains* J. Am. Soc. Mass Spectrom. **2001**, *12*, 1.

(66) Li, Z.; Yalcin, T.; Cassady, C. J. *C-terminal amino acid residue loss for deprotonated peptide ions containing glutamic acid, aspartic acid, or serine residues at the C-terminus* J. Mass Spectrom. **2006**, *41*, 939.

(67) Men, L.; Wang, Y. *Fragmentation of the deprotonated ions of peptides containing cysteine, cysteine sulfinic acid, cysteine sulfonic acid, aspartic acid, and glutamic acid* Rapid Commun. Mass Spectrom. **2006**, *20*, 777.

(68) Pu, D.; Cassady, C. J. *Negative ion dissociation of peptides containing hydroxyl side chains* Rapid Commun. Mass Spectrom. **2008**, *22*, 91.

(69) Pu, D.; Clipston, N. L.; Cassady, C. J. *A comparison of positive and negative ion collision-induced dissociation for model heptapeptides with one basic residue* J. Mass Spectrom. **2010**, *45*, 297.

(70) Waugh, R. J.; Bowie, J. H.; Hayes, R. N. *Collision induced dissociations of deprotonated peptides: Dipeptides containing phenylalanine, tyrosine, histidine and tryptophan* Int. J. Mass Spectrom. **1991**, *107*, 333.

(71) Kulik, W.; Heerma, W. *Fast atom bombardment tandem mass spectrometry for amino acid sequence determination in tripeptides* Biol. Mass Spectrom. **1989**, *18*, 910.

(72) Van Setten, D.; Kulik, W.; Heerma, W. *Isomeric tripeptides: A study on structure-spectrum relationship* Biol. Mass Spectrom. **1990**, *19*, 475.

(73) Harrison, A. G.; Siu, K. W. M.; El Aribi, H. *Amide bond cleavage in deprotonated tripeptides: a newly discovered pathway to " b(2) ions* Rapid Comm. Mass Spectrom. **2003**, *17*, 869.

(74) Bowie, J. H.; Brinkworth, C. S.; Dua, S. *Collision-induced fragmentations of the (M-H)$^-$ parent anions of underivatized peptides: An aid to structure determination and some unusual negative ion cleavages* Mass Spectrom. Rev. **2002**, *21*, 87.

(75) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Fourier transform ion cyclotron resonance mass spectrometry: a primer* Mass Spectrom. Rev. **1998**, *17*, 1.

(76) Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Graham Cooks, R. *The Orbitrap: a new mass spectrometer* J. Mass Spectrom. **2005**, *40*, 430.

(77) Laskin, J.; Futrell, J. H. *Collisional activation of peptide ions in FT-ICR mass spectrometry* Mass Spectrom. Rev. **2003**, *22*, 158.

(78) Cooper, H. J.; Håkansson, K.; Marshall, A. G. *The role of electron capture dissociation in biomolecular analysis* Mass Spectrom. Rev. **2005**, *24*, 201.

(79) Polfer, N. C.; Oomens, J. *Reaction products in mass spectrometry elucidated with infrared spectroscopy* Phys. Chem. Chem. Phys. **2007**, *9*, 3804.

(80) Polfer, N. C. *Infrared multiple photon dissociation spectroscopy of trapped ions* Chem. Soc. Rev. **2011**, *40*, 2211.

(81) Eyler, J. R. *Infrared multiple photon dissociation spectroscopy of ions in Penning traps* Mass Spectrom. Rev. **2009**, *28*, 448.

Chapter 2.

Experimental aspects

As outlined in the previous chapter, the molecular structures of gaseous peptides and their fragments are analyzed based on their infrared spectra. In order to record IR spectra for mass-selected peptide ions, we apply a combination of tandem mass spectrometry and infrared laser spectroscopy. Background of the experimental methods as well as details of the specific apparatus applied in this thesis are discussed in this chapter.

2.1. Fourier-transform ion cyclotron mass spectrometry (FTICR-MS)

Mass spectrometry comprises a range of techniques that accurately measure the masses of charged particles, atoms and molecules. Common to all mass spectrometers is the ionization of the sample followed by separation and measurement of the mass to charge ratio (m/z) of the produced charged particles. Charged particles from multi-component sample are separated according to their mass to charge ratio by altering their trajectories as a consequence of interaction with electric or magnetic fields (e.g., ions of the same mass to charge ratio will undergo the same deflection in a magnetic field as discovered by Wilhelm Wien in 1899).

The first mass spectrometers were experimental setups used for fundamental studies of matter. The study of gas discharges in the mid-19th century revealed anode and cathode rays as positive ions and electrons, respectively, where, if in magnetic or electric fields, particles were deflected unequally as a consequence of their mass difference. This led to the first discovery of non-radioactive isotopes, in particular ^{20}Ne and ^{22}Ne in 1913 by J.J. Thompson[1](Figure 2.1). Later, in 1919, Francis Aston constructed the first velocity focusing mass spectrograph with mass resolving power of 130, and in 1937 with resolving power of 2000. Francis Aston was awarded the Nobel Prize in Chemistry in 1922.

Further advances by Lawrence, who studied particle acceleration, led to the invention of a cyclotron (Figure 2.2)[2], a particle accelerator where charged particles are first injected into a static magnetic field and then accelerated into spiral trajectories by rapidly varying the electric field (in the radiofrequency (RF) regime). To achieve acceleration, the RF frequency must match the particle's cyclotron frequency in the magnetic field. This phenomenon is called (ion) cyclotron resonance. For this discovery Lawrence received the Nobel Prize in Physics in 1939. Subsequently, Frans Michel Penning developed a device to store ionized

Figure 2.1. A replica of a mass spectrometer used by the physicist J.J. Thompson in the 1910s; *Credit: Creative Commons | Jeff Dahl*

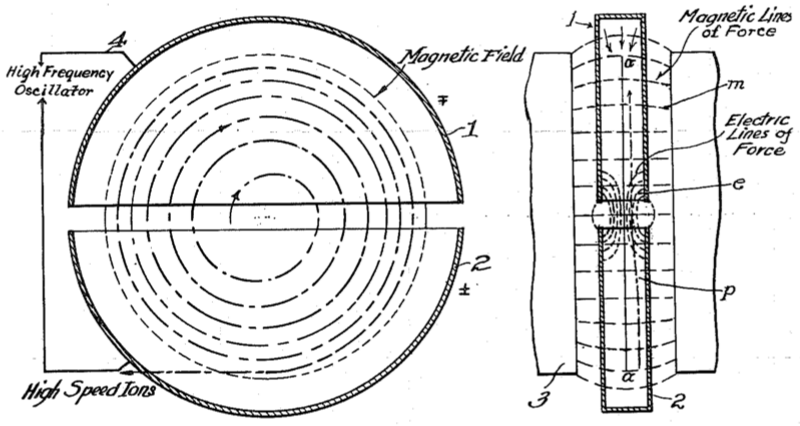


Figure 2.2. Cyclotron. Diagram of cyclotron operation from Lawrence's 1934 patent. The "D" shaped electrodes are enclosed in a flat vacuum chamber, which is installed in a narrow gap between the two poles of a large magnet. (Patent US 1948384 A, 1934.)

species, which was later referred to as the Penning trap. Charged particles are stored by a combination of static electric fields to confine the particles axially and a magnetic field to confine particles radially. In the 1950s, Sommer and co-workers first incorporated the principle of ion cyclotron resonance and the idea of the Penning trap into a mass spectrometer called the omegatron (Figure 2.3).The concept of cyclotron resonance was applied to determine the charge-to mass ratio of the proton[3]. Ion cyclotron resonance, as applied in the omegatron, also forms the basis for modern Fourier-transform ion cyclotron resonance mass spectrometry.

The first FTICR mass spectrometer was invented by A.G. Marshal and M.B. Comisarow[4,5]. The inspiration was found in the ongoing advances in ion cyclotron resonance achievements, as well as in developments of Fourier transform nuclear magnetic resonance (FT-NMR) spectroscopy[3]. Since its discovery, FTICR-MS has emerged as one of the most powerful MS techniques to date, owing in particular to its ultra-high mass resolution[6], wide mass range detection and long ion trapping times.

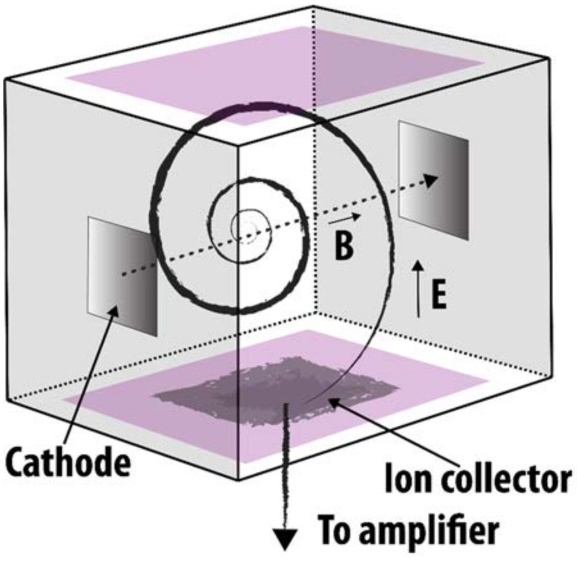


Figure 2.3. Omegatron. In the omegatron, the ions move in circular arcs under the influence of the applied high-frequency electric field E (RF) and the constant magnetic field B. Ions whose cyclotron frequencies coincide with the frequency ω of the field E move in a spiral and reach the ion collector. The mass of these ions satisfies the equation $m/q=a\cdot B/\omega$ where a is a constant for the instrument. In the original design, ions were formed as a result of ionization by a fine electron beam passing through the center of the analyzer along the magnetic field axis.

2.1.1. FTICR theory

The basic principles of FTICR MS can be understood from textbook physics, in particular from Newton's laws of motion and the electric and magnetic forces acting on a charged particle in uniform electric and magnetic fields. Ion cyclotron frequency, radius, velocity, and energy can be derived directly as a function of ion mass, ion charge, and magnetic field strength. Due to the Lorentz force, ions in a magnetic field describe circular orbits around the magnetic field axis. Furthermore, this cyclotron motion of a distribution of ions can be made coherent (which is necessary for the ions to become observable in FTICR-MS) by the application of a temporally varying (RF) electric field at the same frequency as the ion cyclotron frequency. Following this excitation, a free-induction decay signal is recorded from the orbiting ions as a time-domain oscillating "image" current on two (theoretically infinite) opposed parallel electrodes. A frequency-domain spectrum is obtained by Fourier transformation of the digitized time-domain ICR signal. The frequency spectrum can then be converted to a mass spectrum using the ion cyclotron frequency relation $\omega = qB/m$.

Ion motion in a static magnetic field. An ion moving in a spatially uniform magnetic field is subject to the Lorentz force (Figure 2.4) as given by equation 2.1.

$$\vec{F_l} = m \cdot \vec{a} = m\frac{d\vec{v}}{dt} = q\vec{v}\,x\overrightarrow{B} \qquad 2.1$$

In this equation m is mass of the ion, $\vec{a}$ is acceleration, q is the ion's charge, $\vec{v}$ is the velocity and $\overrightarrow{B}$ is the magnetic field. The cross product of $\vec{v}$ and $\overrightarrow{B}$ defines the direction of the Lorentz force to be perpendicular to the plane defined by $\vec{v}$ and $\overrightarrow{B}$ and bends the ion trajectory into a circular orbit (Figure 2.4). The Lorentz force must provide the centrifugal force for the ion to experience this circular motion. If the magnetic field orientation is along the z axis, the velocity of the ion is in the xy plane (v_{xy}) and the angular acceleration (dv/dt) equals v_{xy}^2/r. Equation 2.1 then becomes

$$qv_{xy}B = m\frac{v_{xy}^2}{r} \qquad 2.2$$

If we now substitute v_{xy} in Equation 2.2 using the angular velocity $\omega = v_{xy}/r$, we can derive the equation for the "unperturbed" cyclotron frequency

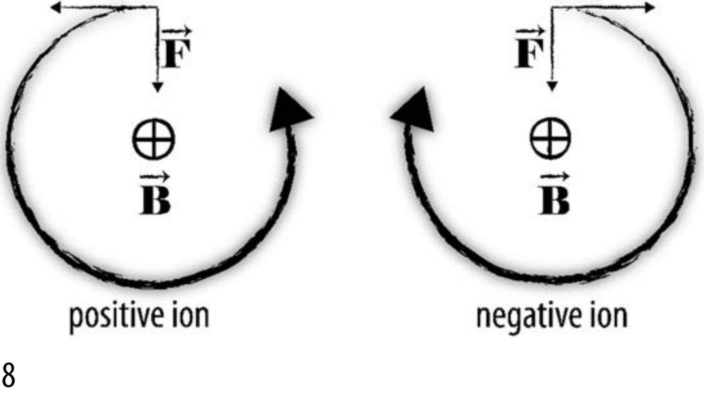


Figure 2.4 The Lorentz force experienced by an ion in a uniform magnetic field.

$$\omega = \frac{qB}{m} \tag{2.3}$$

Writing Equation 2.3 in the terms of frequency instead of angular velocity, one obtains the ion cyclotron frequency in *Hz* (Hertz), where B is in T (Tesla), m is in kg (1 *amu* =1.66 $\cdot 10^{-27}$kg) and q is in C (1 elementary charge *e* equals 1.6 $\cdot 10^{-19}$ C)

$$\nu_c = \frac{qB}{2\pi m} \tag{2.4}$$

From the Equation 2.4, one notices that all ions of same mass to charge ratio have the same ICR frequency, which is moreover independent of their velocity. This makes ICR especially useful for mass spectrometry as not all ions of the same mass to charge ratio (m/z) are required to enter the cell with the same velocity. Hence in contrast to time-of-flight and sector mass spectrometers, ion "focusing" is not essential for the precise determination of m/q. Secondly, for an ion in equilibrium with its surroundings at temperature T(K), the thermal velocity v_{xy} is approximated as $v_{xy} = \sqrt{2kT/m}$. If we incorporate this in Equation 2.2, we obtain an expression for the radii of the ion's cyclotron motion as:

$$r_c = \frac{mv_{xy}}{qB} = \frac{\sqrt{2mkT}}{qB} \tag{2.5}$$

In the 4.7 T uniform magnetic field of our instrument, a singly charged ion at room temperature of 10 amu has radius of cyclotron motion 0.015 mm, 100 amu 0.0487 mm, 1000 amu 0.15 mm, etc. Thus, ions of a wide mass range are confined by the magnetic field to relatively small radii, which is important for their detection as will be explained in the upcoming paragraphs.

Axial confinement of ions. The magnetic field only traps ion motion in two dimensions and ions are free to escape along the z-axis. To avoid such an escape, a small potential, V_{trap}, (usually on the order of volts to tens of volts) is applied using trapping electrodes (see Figure 2.5). The combination of magnetic and electric fields makes the ICR cell a three-dimensional penning ion-trap. This introduces a perturbation to the idealized cyclotron motion of ions; the motion of a charged particle in an FTICR, as it is a penning trap, consists of three different contributions. First there is the axial motion parallel to the magnetic field that results from the applied trapping potential. Then there are the cyclotron motion and magnetron motion, which are both perpendicular to the magnetic field. Trapping motion and cyclotron motion are not coupled, as they are a consequence of only the trapping potential and the magnetic field, respectively. However, the magnetic field in combination with the electric field introduces a third motion of the ions, which is called magnetron motion. Quantitatively these frequencies can be represented with Equations 2.6-2.9 as (detailed derivation of these equations can be found in the review by Marshall and coworkers[4]) :

Axial Trapping frequency

$$\omega_z = \sqrt{2qV_{trap}\alpha/md^2} \qquad 2.6$$

"Unperturbed" cyclotron frequency

$$\omega_c = \frac{qB}{m} \qquad 2.7$$

Magnetron frequency

$$\omega_- = \frac{\omega_c}{2} - \sqrt{(\frac{\omega_c}{2})^2 - \frac{{\omega_z}^2}{2}} \qquad 2.8$$

"Reduced" cyclotron frequency

$$\omega_+ = \frac{\omega_c}{2} + \sqrt{(\frac{\omega_c}{2})^2 - \frac{{\omega_z}^2}{2}} \qquad 2.9$$

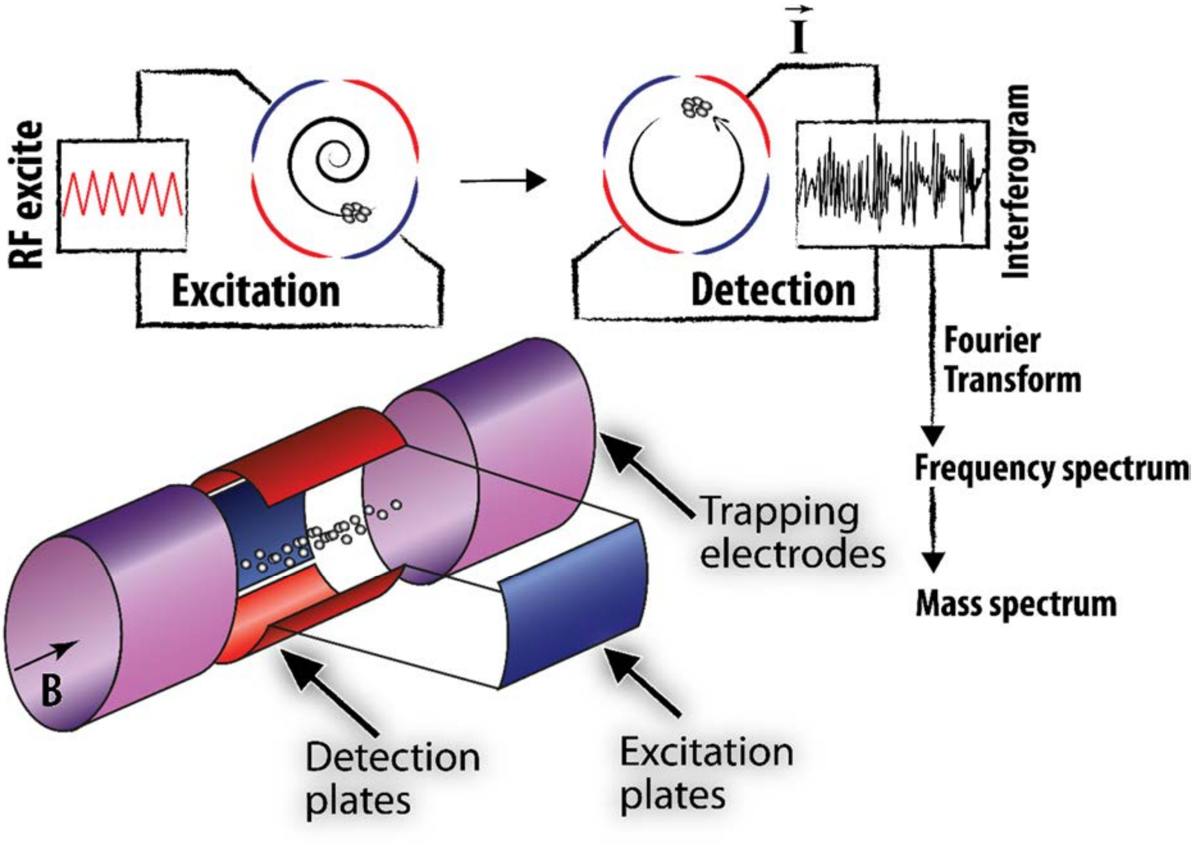


Figure 2.5. FTICR. Representation of an open-ended cylindrical ICR cell which is oriented such that the principal cell axis is aligned with the magnetic field $\vec{B}$. The trapping electrodes are two cylinders at the ends (violet). The center cylinder is segmented into four electrodes (red and blue) that function as excitation and detection plates. The radius of an ion`s cyclotron motion when first trapped is usually small. For ions to be detected, they first need to be excited to coherent motion and larger radii by applying an RF field via the excitation plates. An alternating image current is recorded as ion packets pass the detection plates, generating an interferogram (transient). Using a Fourier transformation this can be translated into a frequency spectrum. Using the ion cyclotron frequency relation, the frequency spectrum can be translated into a mass spectrum.

In Equation 2.6, α is a constant for a particular ICR instrument and depends on the cell geometry while d is the distance between the trapping electrodes. We see that the trapping potential does not depend on the radius r, indicating that the motion is independent of an ion's position inside the trap. The sum of the magnetron and "reduced" cyclotron frequency equals the "unperturbed" cyclotron frequency. The overall motion of ions is a superposition of these three motions and frequencies are related to each other by[7]

$$\omega_c^2 = \omega_+^2 + \omega_-^2 + \omega_z^2 \qquad 2.10$$

Schematically, all motions are presented in Figure 2.6. Magnetron frequencies are on the order 1-100 Hz and are much lower than cyclotron frequencies, which fall in the kHz to MHz range. Further analysis of Equations 2.6-2.9 shows that the magnetron and reduced cyclotron frequencies converge to a common value $\omega_+ = \omega_- = \omega_c/2 = qB/2m$ at the so called "critical" m/q.[8] When $(\frac{\omega_c}{2})^2 = \frac{\omega_z^2}{2}$,

$$(m/q)_{critical} = B^2a^2/4V_{trap}\,\alpha \qquad 2.11$$

For $/q > (m/q)_{critical}$, the ion cyclotron motion is no longer stable and the ion escapes from the trap. From here it is also clear that by lowering the trapping potential, one can increase the upper mass limit.

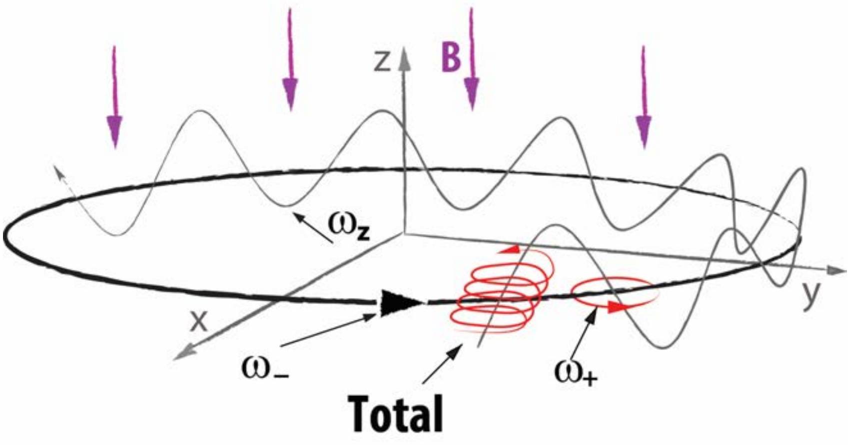


Figure 2.6. All three types of ion motion in an FTICR. Magnetron motion is the motion with a large radius and small angular frequency of ω_-. On top of this, there is an axial motion with a frequency ω_z in the axial direction. Additionally, there is a fast spinning reduced cyclotron motion with frequency ω_+ on top. The overall motion is a superposition of these three components.

Event sequence and ion detection. The most basic FTMS experiment consists of four events: quench, ion formation, ion excitation and detection as illustrated in Figure 2.7. As will be shown later in this thesis, one MS experiment can have many events in its time sequence. Quenching or ion ejection is used to empty the ICR cell of all ions accumulated in the previous event. This can be accomplished, for example, by applying an asymmetric voltage (typically -10V and +10V respectively) on the trapping electrodes, resulting in axial ejection. Then in the ionization event, a new set of ions is formed. They can be formed inside

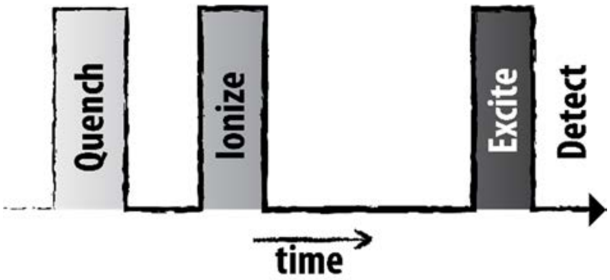


Figure 2.7. Simple FTICR-MS sequence illustrating the basic steps in recording a mass spectrum. The principal functions of mass analysis and ion detection occur in the same space (ICR cell) but are spread in time.

the cell, or by injection from an external ionization source via, for example, a set of RF ion guides[9] and electrostatic lenses[10].

The next step in this sequence is excitation, which refers to the increasing of the ions' cyclotron radii and inducing coherent ICR motion. The radius of an ion`s cyclotron motion when first trapped inside the ICR cell is usually small when compared to the dimension of the cell. It is important to recognize that ion cyclotron motion itself is not so useful; a small cyclotron radius does not allow an image current to be detected on the detector electrodes (Figure 2.5). Moreover, the phase of orbital motion for each ion is random, so that any current induced on the detector plates would be counterbalanced and the overall signal would average to zero. This necessitates coherent excitation to larger (and detectable) orbital radii, which is achieved by applying a electric RF field oscillating at or near the cyclotron frequency of the ions.

From Equation 2.5 it can be seen that the radius of ion cyclotron motion depends on the velocity and therefore on the kinetic energy. Rearranging this equation, we can write the expression for the kinetic energy, K_{xy}, of the ion corresponding to its motion in the xy plane as

$$K_{xy} = mv_{xy}^2/2 = q^2B^2r^2/2m \qquad 2.12$$

When applying the RF field, ion motion is still described by the Lorentz force equation, but now there is the additional presence of an electric field, which can be described with

$$\vec{E}(t) = E_0\cos(\omega t)\vec{j} \qquad 2.13$$

Here, j is the unit vector and E_0 is generated by applying $+V_0(\cos(wt))$ and $-V_0(\cos(wt))$ on two opposite parallel electrodes separated by distance, d:

$$E_0 = 2\,V_0(\cos(wt))/d = V_{p-p}/d \qquad 2.14$$

The uniform electric field E_0 can be separated in two oppositely rotating components E_L and E_R. The RF component of E_0 rotating in the same direction and with the same frequency ("resonant with") as the ion of interest is pushing the ion outward in its orbit (Figure 2.5). All ions of the same mass-to-charge ratio are excited coherently, meaning that they are grouped as tightly after excitation as they were initially. As the radius of cyclotron ion motion increases, the ion packets pass closer and closer to the detector electrodes.

Following RF excitation, the oscillating image current induced by the coherently orbiting ions is picked up by a separate set of electrodes. The time domain signal that is recorded is a free induction decay (FID), also referred to as transient or interferogram (Figure 2.5). It consists of a superposition of sine waves produced by each of the ionic species in the population. Fourier transformation of this signal gives the frequency components contained in the transient. This frequency spectrum can then be converted to a mass spectrum using Equation 2.3.

The power absorption, P, by the ion from the RF field can be determined as the amount of work done by the electric force, $F \cdot s$, in time s, $P = F \cdot s/t = F \cdot v_{xy} = qE \cdot v_{xy}$. Both Comisarow[11] and Dunbar[12] solved this equation with the assumption that the frequency of de-energizing collisions is much lower than ω_c (at typical ICR pressures <10^{-8} mbar, the mean free path amounts to several km or more, so that collisional damping is negligible), and that ω_c is approximately the same as the excitation frequency:

$$P_{t_{rf}} = q^2 E_0^2 t_{rf}/4m \qquad 2.15$$

where t_{rf} is the time of RF excitation. The kinetic energy acquired by the ions from the RF field is then calculated as the integral of power absorption over time as $K = \int_0^t P(t_{rf})dt_{rf}$. Thus, the post-excitation kinetic energy of the ions can be expressed as

$$K = q^2 {E_0}^2 {t_{rf}}^2/8m \qquad 2.16$$

In combination with Eq. 2.12 this leads to the expression for the radius of the ion as a function of the RF amplitude and irradiation time:

$$r = E_0 t_{rf}/2B \qquad 2.17$$

In Eq. 2.17 the post-excitation cyclotron radius is independent of ion mass and charge (m and q). Hence, all ions in a given mass range can be excited to the same ICR orbital radius and detected simultaneously. Usually, excitation electrodes are separated by a few centimeters and the voltage applied is on the order of tens of volts. Post-excitation radii are usually on the order of centimeters, roughly defining the inner dimensions of the ICR cell. To excite all ions within a certain mass range, the RF excitation field is usually swept over the corresponding frequency range.

High mass resolution is one of the major advantages of FTICR in comparison with other types of mass spectrometers. In mass spectrometry, resolution measures the ability to differentiate two ions of slightly different mass-to-charge ratios Δm/z. Resolution can be defined as the full width at half-maximum (FWHM) of a peak, $\Delta\omega_{50\%}$ and $\Delta m_{50\%}$, for frequency and mass peaks, respectively. **Resolving power is** defined as $m/\Delta m_{50\%}$ and it can be expressed as

$$m/\Delta m_{50\%} = q\ B_0/m\Delta\omega_{50\%} \qquad 2.18$$

The mass resolution increases proportionally to the length of the transient time, with increasing number of data points collected, as well as with increasing magnetic field intensity. Mass resolution decreases with increasing pressure in the cell, which leads to quicker dephasing of the FID.[4,13]

The strength and accuracy of FTICR MS can be illustrated by an experiment by Marshall et al. from 2001, where baseline resolution of two peptide isobars was achieved with a mass resolving power of 3300000. The elemental composition of the peptides differed by N_4O versus S_2H_8 (Δm = 0.00045 Da), less than the mass of a single electron (0.00055 Da). Developments used in these experiments included a 9.4 T magnetic field, a large diameter penning trap (4 inch) as well as a low ion density to reduce space-charge effects.[6,14]

2.1.2. FTICR mass spectrometer

We describe the design of an FTICR mass spectrometer as consisting of four main elements: the magnet, analyzer ICR cell, vacuum system and data system (the vacuum and data systems will be subject of Chapter 2.3.1).

The magnet can be either an electromagnet or a superconducting magnet, with magnetic field strengths demonstrated from a few Tesla up to 20 T, respectively [15]. Superconducting magnets are often used to reach high and homogeneous magnetic fields. Superconducting magnet coils are made of materials (usually alloys of metals such as niobium-titanium $(Nb,Ti)_3Sn$, vanadium-gallium (V_3Ga) etc.) that conduct electric current without resistance. To reach superconductivity, they must be cooled to cryogenic temperatures during operation. This is achieved by embedding the coils inside a cryostat that contains three level cooling – first, the system is embedded in a vacuum cylinder jacket for reducing heat transport from room temperature, second is a liquid nitrogen shield, and third, the liquid helium vessel that provides the critical temperature (e.g. the temperature below which the coil material becomes superconductive) for the used material. Cryogenic temperatures must continuously be maintained to prevent quenching of the magnet. A quench is an abnormal termination of magnet operation that occurs when a part of the superconducting coil enters the normal (resistive) state. The amount of energy stored inside the superconducting magnet is enormous, between a few hundred thousand to millions of Joules! Upon quenching, this energy is converted into heat, which can damage the superconducting coil to the point of no return and moreover results in the rapid evaporation of the remaining coolants. When the coil is charged, the continuous current generates the magnetic field. In its superconducting state, the wire can conduct much larger electric currents than ordinary wire (for example, the 4.7 T superconductive magnet that will be described in upcoming paragraphs operates with a current of 70 A), creating intense magnetic fields. The ICR cell is set in the center of the bore of the magnet, where the field is maximum and moreover tuned for the highest homogeneity (typically on the order of 10 ppm).

ICR analyzer cell. Geometries for ICR cells can be divided into two basic groups according to their shape: the cubic cell (where all six electrodes are flat and assembled in a cube) and the open-ended cylindrical cell. A schematic overview and real image of a cylindrical, open-ended ICR cell are presented in Figure 2.5 and Figure 2.8, respectively. It features two trapping electrodes and two sets of opposite detection and excitation electrodes. The role of each is explained in detail in the theoretical description of the FTICR earlier in this chapter.

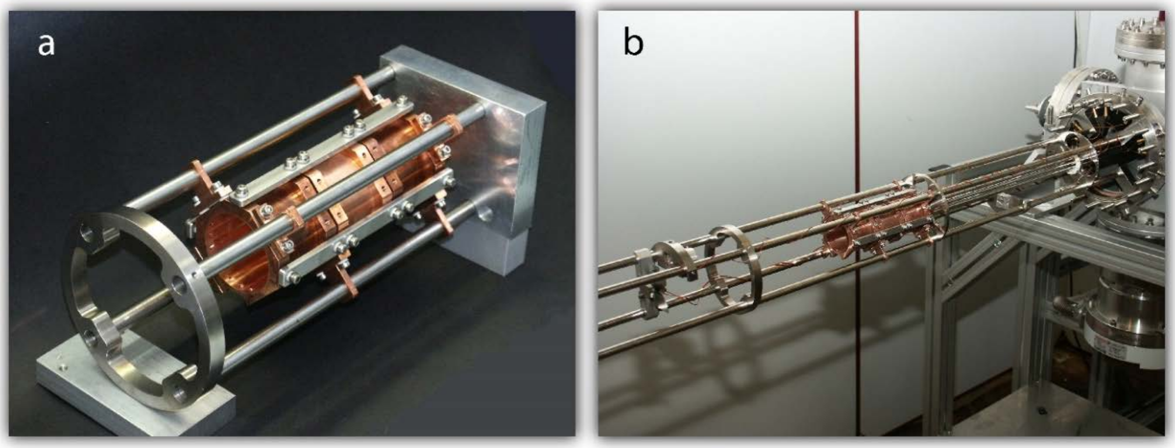


Figure 2.8. The ICR cell is made of copper plates arranged cylindrically and set inside the vacuum manifold (a). The octopole ion guide is used to transfer the ions inside the magnetic field of ICR cell (b). This ICR cell is part of the FTICR mass spectrometer that was used in performing the experiments in this thesis.

Various ion optics are used to transport ions from the ionization source to the ICR cell. These elements can vary significantly from instrument to instrument and can consist of sets of electrostatic or RF ion guides, linear RF traps that are used for ion storage/reaction/fragmentation prior the ICR MS analysis. High vacuum (10^{-8}-10^{-10} Torr) is required inside the ICR cell for achieving high resolution of an instrument (higher pressures would increase the number of ion-neutral collisions inside the ICR cell , shortening the length of the transient and leading to lower mass resolution).

Ion formation. Every mass spectrometer uses a set of magnetic and electric fields to manipulate and analyse charged particles (ions). Thus, an analyte compound must be charged or ionized to be observed in the mass spectrometer. Moreover, those ions must be brought into the gas phase and then into the high-vacuum region of the mass spectrometer. This is almost trivial for samples which are gaseous or easily volatilized. However, many compounds are thermally labile and decompose when heated. These types of samples necessitate either desorption or desolvation methods. Even if ionization and desorption/desolvation are predominantly separate processes, the expression "ionization method" is customarily used to address both ionization and desorption/desolvation approach.

The choice of ionization method for a given experiment depends on the physical and chemical properties of the sample and the type of information desired from the analysis. ‘Soft ionization' methods such as field or laser desorption and electrospray ionization likely produce mass spectra with negligible ion fragmentation and are found especially convenient in ionizing molecules or molecular complexes of biological interest. Several “soft ionization” methods have been demonstrated to date[16-18]. However, only electrospray ionization will be described in more detail here, as it has been used for the production of ions described in this thesis.

2.1.3. Electrospray ionization (ESI)

Design and simplicity of operation, diverse capabilities and a wide range of applications have made electrospray a very popular and widely used ionization technique. ESI can be used to produce ions of polymers, nucleic acids, peptides, proteins, carbohydrates and many other compounds. The sample can be a cation or an anion and the produced ions can be singly (+/-1) charged, or multiply charged (+/-n charges). Positive ions are typically generated by proton or metal ion addition, while negative ions are typically generated by deprotonation or sometimes halide attachment. The sample can be a simple salt solution or a very complex protein mixture.

When forcing a continuous flow of a conducting liquid through a metal capillary, maintained at higher voltage relative to the counter electrode, electrospray (ES) occurs. Such conditions result in spraying and aerosol formation, a phenomenon that was first detailed and photographed by John Zeleny at the beginning of the 20th century.[19] Further experiments leading to the development of electrospray ionization as we know it today were performed by M. Dole in the 1960s. Actually, he discovered the possibilities of electrospray by a lucky coincidence while visiting a car manufacturer where cars were painted using electrospray (electrospray was able to produce small charged particles of paint to be attached on the metal surface of a car). Using this as an idea, Dole and co-workers were able to develop an apparatus that produced charged polystyrene molecules in the gas phase with molecular weights in the kilo-Dalton range.[20] However, only in the mid-1980s this method was first implemented in mass spectrometry by John Fenn[17,21,22]. For this development he was awarded the Nobel Prize in chemistry in 2002.

In Figure 2.9, the basic electrospray mechanism producing positive ions is shown. The sample solution enters the atmospheric pressure part of the electrospray chamber through a conductive capillary (“needle”) having a diameter on the order of 0.1 mm. The solution is supplied by a mechanical syringe pump. For regular ESI, the sample concentrations may vary from μmol/l to mmol/l. The flow rate is usually tens to hundreds of μl/hour. The ESI needle is held at high voltage, varying from ~1.5-6 kV depending on the parameters of the instrument or experimental needs. The polarity of the applied voltage defines the polarity of the generated ions. Typical solvents used in ESI are methanol, acetonitrile and water,

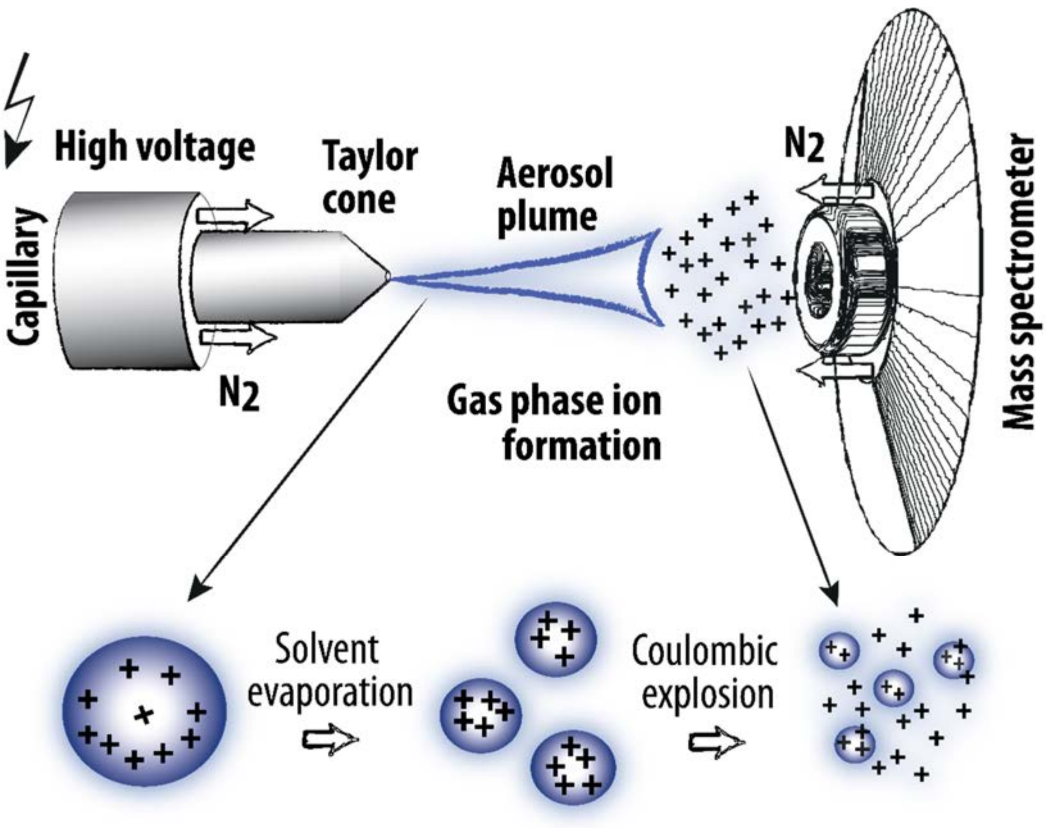


Figure 2.9. ESI The analyte solution flows through the electrospray capillary held at high voltage. A large potential difference, usually few kV (with respect to the counter electrode), drives the spraying of charged droplets from the needle having charge of the same polarity as the charge on the needle (aerosol plume). As the droplets traverse the space between the needle tip and the cone, solvent evaporation occurs. When the repulsion overcomes the surface tension, droplets burst into smaller droplets in so called Coulomb fission.

usually in a solvent mixture. Generally speaking, solvents having a high surface tension are not a good choice and it is better to use them in a mixture (water has high surface tension and spraying from pure water is more challenging). The desolvation and nebulising gas (typically N_2, normally heated to approximately 100°C) aids the stabilization of the spray and facilitates the evaporation of the solvent molecules from the charged droplets. The electric field at the tip of ESI capillary can be estimated to be[23]

$$E_c = \frac{2V_c}{r_c ln(4d/r_c)} \qquad 2.19$$

where V_c is the applied potential, r_c the capillary outer radius and d is the distance from the capillary tip to counter-electrode (or entrance to mass spectrometer as presented in Figure 2.9). For the standard values of the dimensions in the Eq. 2. (V_c~2 kV, r_c~ 0.5 mm and d ~0.02m) we can calculate that the electric field at the capillary tip is 1.6×10^6 V/m[24]. Such a high electric field penetrates into the solution and causes the polarization of the solution; the surface of the liquid meniscus is now enriched with positive ions while negative ions move away from the meniscus. This distorts the liquid meniscus existing at the needle tip into a so called Taylor cone[25,26]. If the applied field is sufficiently high and the surface tension is overcome, the liquid surface becomes unstable and emerges in a fine liquid jet. The formed jet is populated with the excess of positive charges and their mutual repulsion causes the jet to break up into small charged droplets.

The droplets as well carry excess positive charge as they travel through the source toward the opposing electrode. At the same time the solvent evaporates and leads to droplet shrinkage and an increased potential at the surface. Droplets shrink until they reach the Rayleigh limit[27]. Again, when the repulsion overcomes the surface tension, the droplet breaks up into smaller droplets in so called Coulomb explosion (or fission). The charge q of the droplet and surface tension γ are related as follows[27]:

$$q = 8\pi\sqrt{\varepsilon_0 \gamma R^3} \qquad 2.20$$

Where R is the droplet radius and ε_0 is the vacuum permittivity. This process repeats in cascades and results in charged (singly or multiply) individual, isolated molecules that enter the mass spectrometer for analysis.

However, the final events in droplet desolvation have been under much debate[28,29] and there are two hypotheses used to explain this phenomenon – the *charge residue* and *ion evaporation* models. In the charge residue model, first proposed by Dole et al.[20], electrosprayed droplets undergo evaporation and fission cycles, eventually leading to progeny droplets. Gas phase ions form after the remaining solvent molecules have evaporated, leaving the bare ion with all the remaining charges that the droplet carried. This model was mainly used to explain highly charged ion generation in ESI of relatively large molecules such as proteins, polymers, etc.[30]

In the ion evaporation model, direct emission of ions from the droplet surface is proposed. After the droplet reaches a certain radius following a series of coulomb fissions, the field strength at the surface assists the desorption of solvated ions.[31] According to direct evidence, both of the proposed mechanisms are valid for different cases. Moreover, it has been shown that in cases where the sample solution contained high concentrations of salts (as small ions) and macromolecular ions (as large ions), both of the mechanisms occur simultaneously.[32] A simplified schematic overview of this combined model is presented in Figure 2.9.

2.2. Infrared light sources – FEL and OPO

When the first laser was reported in 1960, it was described as "a solution looking for a problem."[33] Today, lasers can be found everywhere. Since the construction of first ruby laser by Maiman in 1960[34] (Figure 2.10), there has been a tremendous advance in the field of laser science. Lasers ranging in size from tiny diode lasers to enormous room or field-sized ones (like FELs or x-FELs) are utilized in a variety of applications in every section of modern society, including information and communications processing, entertainment, medicine, industry and science. The laser triggered the photonics revolution as well as delicate experiments in the field of modern quantum physics.

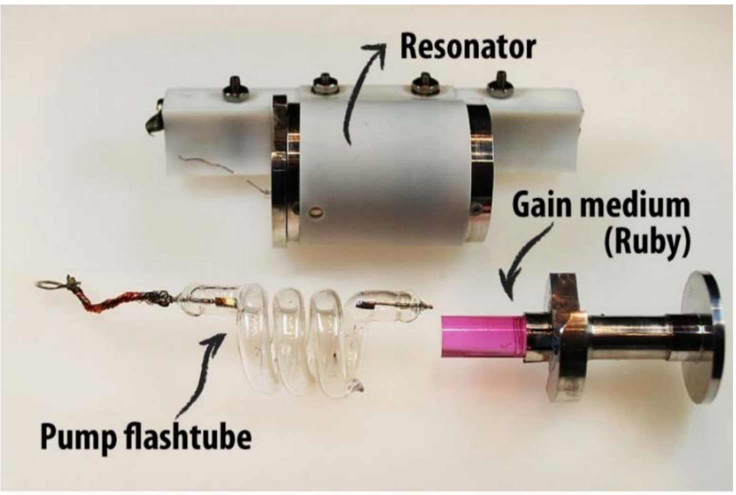


Figure 2.10. Parts of the ruby laser
The first laser was a ruby laser made by T. Maiman at Hughes Research Laboratories on May 16, 1960. A ruby laser consists of a ruby rod that is optically pumped with flash lamp to achieve population inversion. The rod is placed between two mirrors, forming an optical cavity – resonator.

Laser. The word 'laser' is an acronym for light amplification by stimulated emission of radiation. The basic operating principles of a laser involve the quantum nature of matter and stimulated emission for amplification of light, finding its origin in an idea Einstein developed in 1916. Stimulated emission is the process by which an electron in an excited molecular or atomic orbital, interacting with an electromagnetic wave of the appropriate frequency, may relax to a lower energy orbital and emit a photon corresponding to the energy difference. A new photon created in this manner has the same phase, frequency, polarization, and direction of travel as the photons of the incident field (stimulated emission can provide a physical mechanism for optical amplification). Graphical representation of this process is presented in Figure 2.11.

Basic components of a laser include an optical cavity called the optical resonator, a laser gain medium and a pump source to excite the gain medium (see Figure 2.12). The optical resonator consists of at least two mirrors between which the light resonantly bounces. One of the end mirrors is only partially reflective (or holed), so that a small portion of the light is transmitted forming the laser output. The stimulated emission process takes place in the gain medium. Lasers are typically classified by the type of gain medium they employ (gas laser, solid-state laser, dye laser, semiconductor laser, free electron laser, etc.). The gain medium amplifies light of any direction but only the light that is resonated in the cavity

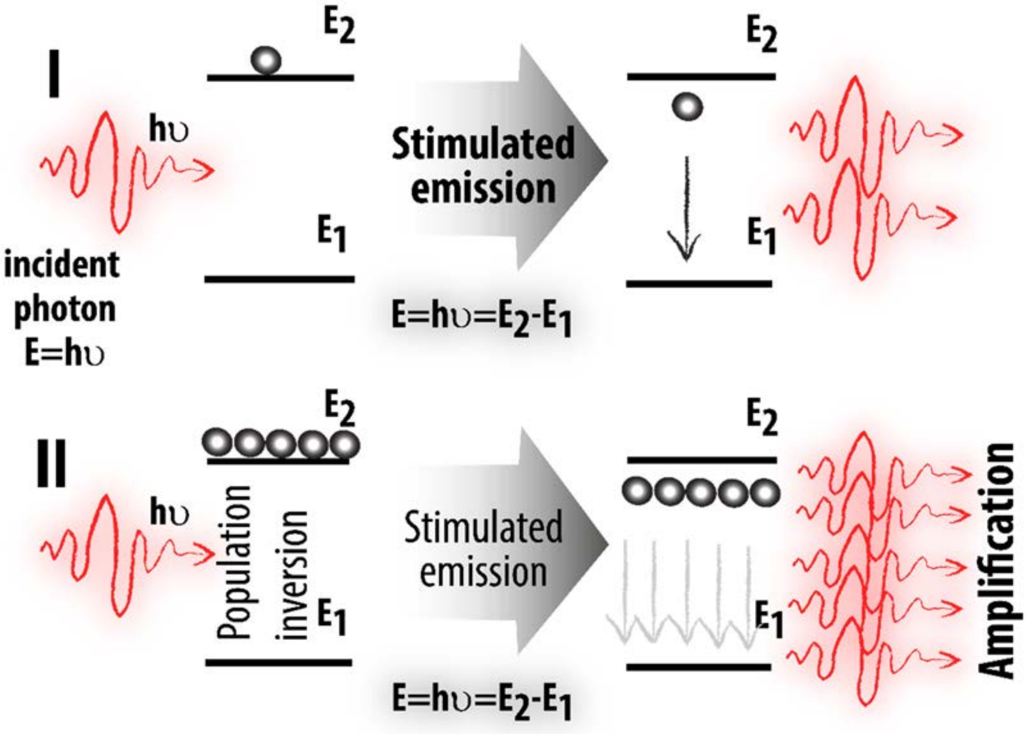


Figure 2.11. Stimulated emission and light amplification. **(I)** When an electron is in an excited orbital (E_2), an incoming photon for which the energy is equal to the energy difference between two orbitals 1 and 2 ($h\nu = E_2 - E_1$) can "stimulate" a transition to that lower level, producing a second photon of the same energy. **(II)** If the population of electrons in the higher energy orbital is larger than that in the lower energy orbital (a condition called "population inversion"), the stage is set for stimulated emission of multiple photons. Population inversion is a precondition for the light amplification which occurs in a laser. The emitted photons have a definite time and phase relation to each other and the light has a high degree of coherence.

(passing more than once through the medium) will be significantly amplified and therefore reach a high intensity. The fact that the photon energy has to match a given energy transition makes the laser monochromatic. Since the amplification process maintains the phase and direction of the light, the laser output is directional and coherent.

For light amplification in the medium to dominate over absorption by the medium, a situation of population inversion (Figure 2.11) is required, in which a majority of the active particles of the gain medium are in the excited state. To achieve population inversion, a pumping process is needed that continuously brings the particles into the upper energy state. An example of a pump source is optical pumping by flash lamps or by another laser.

The characteristics which distinguish lasers from ordinary light sources are monochromaticity (ordinary light sources, such as the incandescent lamp, are a combination of many different wavelengths, while a laser, on the other hand, emits only a very narrow range of wavelengths), directionality (a light bulb, emits light in all directions, while lasers can emit light that diverges only very little with distance) and coherence (which means that the light is in phase spatially and temporally).

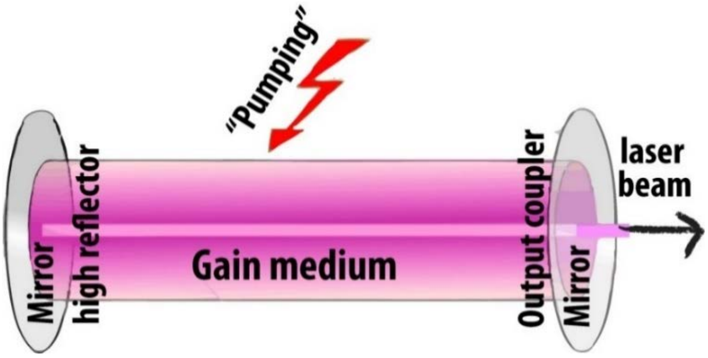


Figure 2.12. The basic laser concept. A laser consists of a gain medium, an energy supplier, and an optical feedback. The gain medium is a material with properties that allows it to amplify the light by stimulated emission. For the gain medium to amplify the light, it needs to be supplied with energy. That is the "pumping" process. Light of a specific wavelength is then amplified, passing through the gain medium with multiple reflections between the mirrors (increasing the power of the output light). The laser beam is coupled out through the semitransparent mirror.

2.2.1. Free electron laser (FEL)

The wavelengths that can be provided by a laser relying on atomic or molecular electronic transitions in the laser gain medium are limited by the electronic structure of the gain medium material. Several researchers realized that this can be circumvented by using a laser based on emission from free electrons.[35-37] In the free electron laser, coherent radiation is produced by unbound electrons traversing in an oscillating trajectory, instead of being defined by the electronic transitions in atoms, molecules or crystals. The free electron laser, as it is today, has the widest frequency range of any laser type, tunable from microwaves, through terahertz radiation and infrared[38], to the visible, ultraviolet, and X-ray region of the spectrum [39]. There are currently no alternative laser sources that feature a comparable combination of tunability, high pulse energy and short pulse duration.

In a free electron laser, radiation is produced by relativistic electrons which "wiggle" under the influence of a magnetic field of periodically changing polarity. Figure 2.13 shows a scheme of a free electron laser (with a cavity; X-FELs are built on the same principle, but since mirrors for X-rays, as their photons fall within the 0.1 to 10 nm range, are not available, they have a different design). Electrons are produced in an electron gun, bunched (the electron bunch is approximately a few picoseconds long) and accelerated (for example by a linear electron accelerator) to relativistic speeds. Relativistic electrons are then injected into an alternating magnetic field generated by two arrays of magnets referred to as an undulator, or wiggler. Here, the electrons experience a Lorentz force inducing oscillations in their trajectories. As a consequence of the (angular) acceleration, radiation is emitted with the frequency of the oscillation period and in the direction of the electron beam. The relativistic wiggling electrons can be regarded as a Hertz dipole and thus emit radiation perpendicular to their direction of oscillation. Due to the high velocity of the electrons, a strong Doppler shift of the emitted radiation to higher frequencies is experienced in the laboratory frame. Radiation is in resonance with the electron if radiation advances by an integer number of periods relative to the electron for each oscillation of the electron in the undulator. This resonance condition provides the wavelength λ of the spontaneously

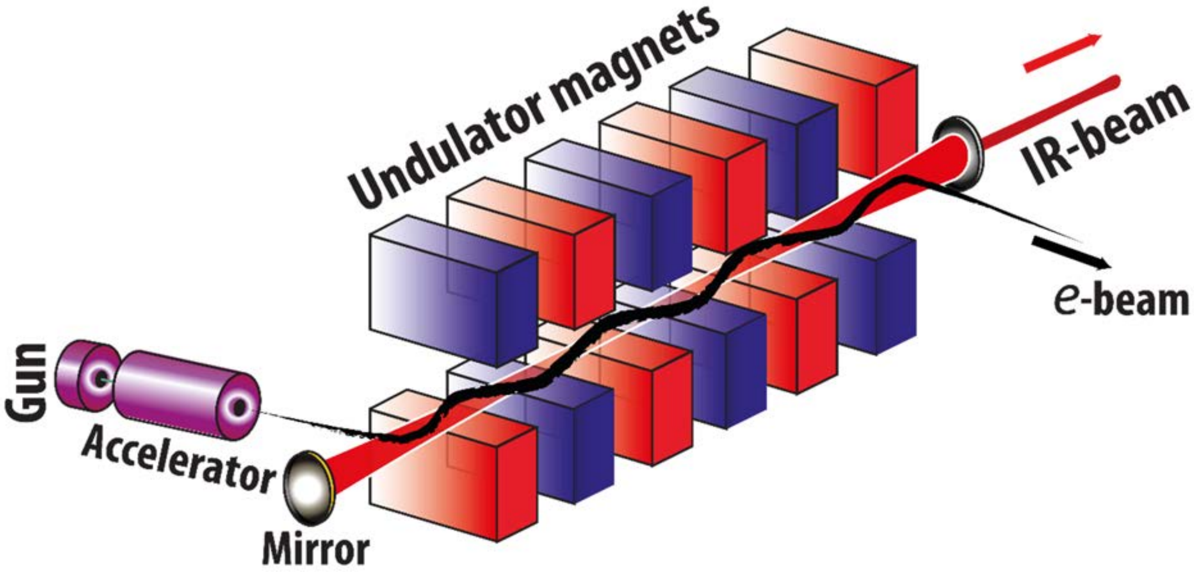


Figure 2.13. FEL Principle Diagram showing the basic ingredients of a free-electron laser: an accelerator to provide a beam of relativistic electrons, an undulator through which the electrons travel inducing oscillation and photon emission, and two mirrors forming an optical resonator in which the radiation is amplified. By changing parameters like the electron energy or the intensity of the magnetic field, different wavelengths can be produced. The relativistic electrons, after finishing their task of producing the light are "dumped".

emitted radiation, depending on the electron energy E, the period of the undulator magnets λ_u and on the magnetic field B, as shown below.

$$\lambda = \frac{\lambda_u}{2\gamma^2}(1 + K^2) \qquad 2.21$$

In this expression, K is a dimensionless parameter accounting for the undulator period λ_u and the magnetic field strength B_u.

$$K = \frac{eB_u\lambda_u}{2\pi m_0 c} \qquad 2.22$$

In Equation 2.22, m_0 is the rest mass of the electron and c is the speed of light. In Equation 2.21, γ is the Lorentz factor, defined as

$$\gamma = 1/\sqrt{1 - v^2/c^2} \qquad 2.23$$

The spatial spread of the electrons inside the undulator after they leave the accelerator is much larger than the emitted wavelength. Thus, the electrons do not emit coherently and the radiation is usually not very intense. In an IR FEL, to force the electrons to emit coherently, the undulator is placed inside an optical resonator composed of two highly reflective mirrors at either side (Figure 2.13). In this configuration, the emitted radiation is stored in the optical cavity. When new electrons enter the cavity, they interact with the optical field in the cavity and experience a so-called ponderomotive force. This force modulates the density of the injected electrons along the longitudinal axis of the undulator, (micro)bunching them to equally spaced slices on the scale of radiation wavelength, leading to coherent emission.

If now we consider only a single electron entering the undulator at the resonance regime, the energy transfer between the electron and the spontaneously emitted light field can occur in both directions. This results in acceleration or deceleration of an electron. However, in a bunch, averaging over all electrons results in a zero net gain as both processes occur to the same extent. Due to already mentioned phase-dependent energy transfer, a density distortion of the electrons on the wavelength scale is developed along the undulator. If the electrons now are introduced into the undulator at an energy which is slightly higher than the resonant energy value, then positive net gain is achieved. Positive gain limitation is at the saturation level, which is approximately 10^6-10^8 times higher than the spontaneous emission. Thus, the free electron laser can generate a very high photon flux.

For X-ray FEL lasers, lack of suitable mirrors in the extreme ultraviolet and X-ray regimes for the cavity construction demands other laser amplification techniques. Therefore, X-ray free electron lasers use long undulators and techniques such as self-amplified stimulated emission (SASE). One of the SASE approaches consists of passing the electron beam through two undulators. In the first undulator, the light of a particular wavelength is produced. Subsequently, both the produced light and the electrons pass through the second undulator which induces the microbunching of initial electron bunch due to the ponderomotive force.[40,41]

The most striking difference of the FEL compared to other lasers is that, in principle, it can be designed to operate at any wavelength. The wavelength of the radiation can be set by changing the electron energy or varying undulator features, such as the magnetic field strength (for example, by changing the distance between the undulator magnets).

2.2.2. Optical parametric oscillator (OPO)

The optical parametric oscillator (OPO) was developed in the 1960's as a tunable mid-infrared light source. Emission from an OPO is the result of the conversion of a pump photon into two output photons with lower frequency, referred to as the signal and idler waves, by a nonlinear crystal.

Energy conservation requires that the sum of the frequencies of the output beams equals that of the input (pump) wave frequency

$$\omega_p = \omega_s + \omega_i \qquad 2.24$$

The momentum conservation needs to be satisfied as well

$$k_p = k_s + k_i \qquad 2.25a$$

$$\nu_p n_p = \nu_s n_s + \nu_i n_i \qquad 2.25b$$

so that new photons can add up to a macroscopic wave only in the case of phase matching described by Equation 2.25. In the Equation 2.25a k is the wavevector (or angular wavenumber), and since $k = 2\pi\nu n/c$ (where n is the refractive index, c is the speed of light and ν the frequency), Equation 2.25a can be expressed as Equation 2.25b. Such matching of phase velocities of waves of different wavelength can be achieved in birefringent crystals, for which the refractive index depends on the polarization and the direction of passing light. To tune the frequencies of the two output beams, the alignment of the optical axis of the crystal relative to the pump beam is varied. Tuning of an OPO laser can also be achieved by controlling the crystal temperature. In an OPO, the optical crystal is placed inside an optical resonator or cavity. The signal and idler beams are produced when the amplified beam leaves through one of the resonator mirrors (the signal beam commonly refers to the higher-frequency beam and the idler is the lower frequency beam).

In order to obtain higher power, after the OPO stage, normally, one or both of the beams is injected into the optical parametric amplification (OPA) stage. This consists of passing the signal or idler wave through another nonlinear optical crystal(s) collinearly with the photons having the same wavelength as the pump photons, allowing the production of more intense beams with the same frequency as the original signal or idler (for more details see Chapter 2.3.3).

2.3. Experimental setup

2.3.1. FTICR mass spectrometer

The combination of an ion trap mass spectrometer with a tunable laser having high output power, such as FELs or an OPO, makes it possible to record infrared (IR) spectra of molecular ions in the gas phase in a process known as infrared multiple photon dissociation (IRMPD). The basic principle relies on monitoring ion fragmentation as a function of laser wavelength, generating an IR spectrum (details of this method will be presented in the following chapter). In this technique, the IR spectrum is not directly recorded by measuring the attenuation of the IR light, but is generated by monitoring a consequence or "action" induced by photon absorption. The collection of IR spectra of ions is useful in a variety of scientific fields: astrochemistry (where spectral characterization of species is used to support the hypothesis of their occurrence in interstellar space)[42-44], biochemistry (where the structure of biomolecules is naturally of key interest)[45-47], and for fundamental chemical problems such as kinetics (studying reaction pathways and intermediates), differentiation of structural isomers or for understanding catalytic processes.[48,49] The first use of a tunable laser combined with ion trapping approaches dates back to the 1970`s, when Beauchamp and co-workers irradiated ions in an FTICR mass spectrometer with the output line of a tunable CO_2 laser.[50]

Here a brief overview of the FTICR mass spectrometer that is coupled to the beamline of the IR free electron laser FELIX and an OPO, as this combination forms the complete experimental setup from which all the results of this thesis were obtained.

FTICR mass spectrometer at FELIX facility is a home-built instrument equipped with a superconducting 4.7 T actively shielded magnet (Cryomagnetics) with a 128 mm diameter horizontal bore (Figure 2.14 and 2.15). Ions are produced using a z-shaped electrospray ionization source (Micromass Z-Spray, Figure 2.16). In the electrospray source, ions are

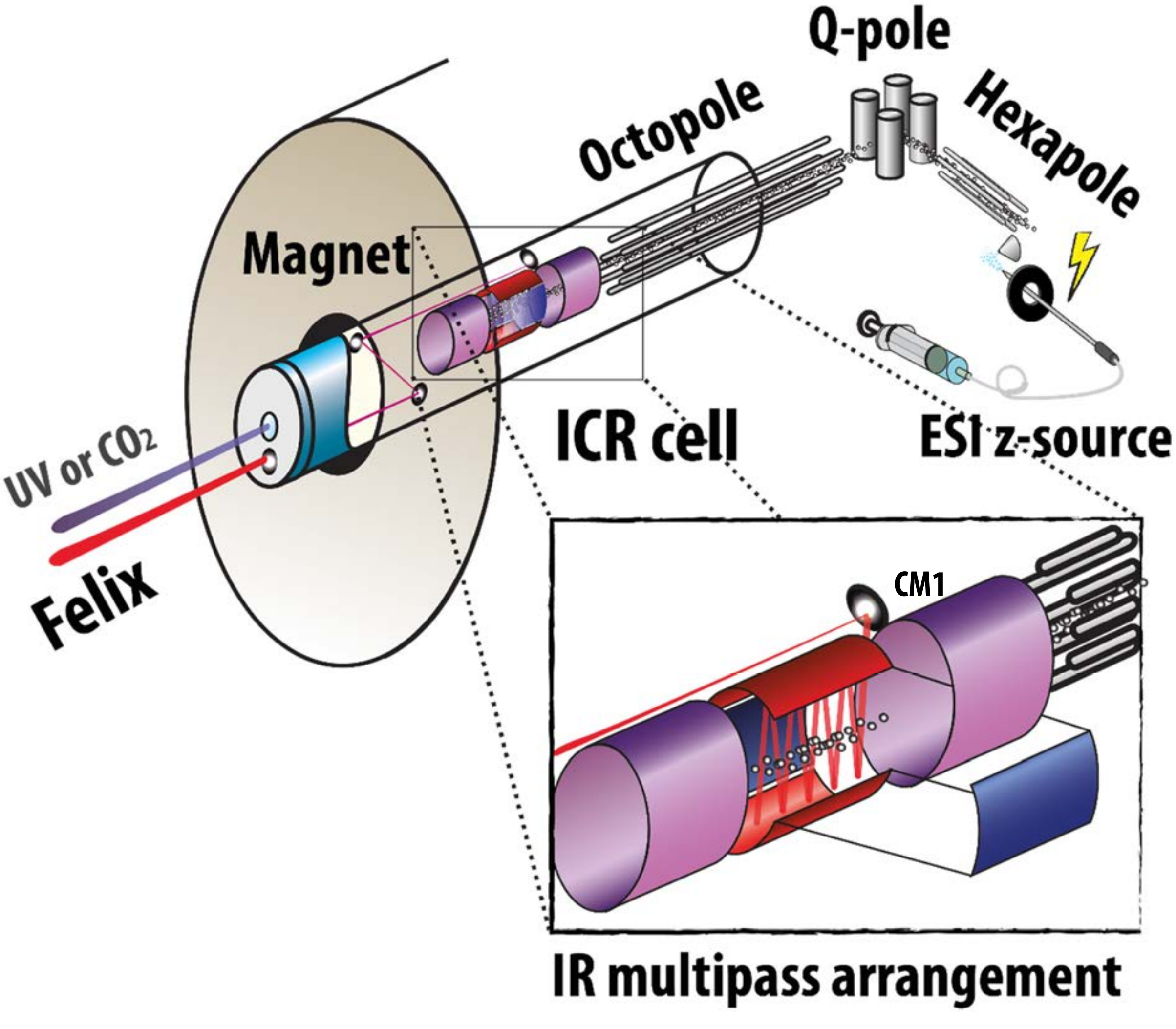


Figure 2.14. FTICR mass spectrometer adapted for IR spectroscopy of ions. Ionization takes place in the Z-electrospray source, where ions are going from solution into the gas phase. Ions are dragged by the potential difference between the capillary and the cone. They are accumulated in the hexapole. Ions trapped in the hexapole can be fragmented by CID (for explanation see 2.3.1 part of this chapter). After the ions leave the hexapole, they are focused using a set of ion lenses (not shown). Subsequently, ions are transferred to the ICR cell via a quadrupole ion deflector and a 1 m long rf octopole ion guide. Once in the ICR cell, the ions can be mass isolated and exposed to FELIX or OPO radiation to record IR spectra. Additionally, a UV or a continuous-wave CO_2 laser can be used to manipulate (e.g. fragment) the ions. The IR laser beam is aligned parallel with, but off-axis, from the centerline of the magnet bore to reach copper mirror CM1, from where it is reflected through the ICR cell (multipass arrangement). This multipass geometry ensures a larger interaction of the trapped ions and the light.

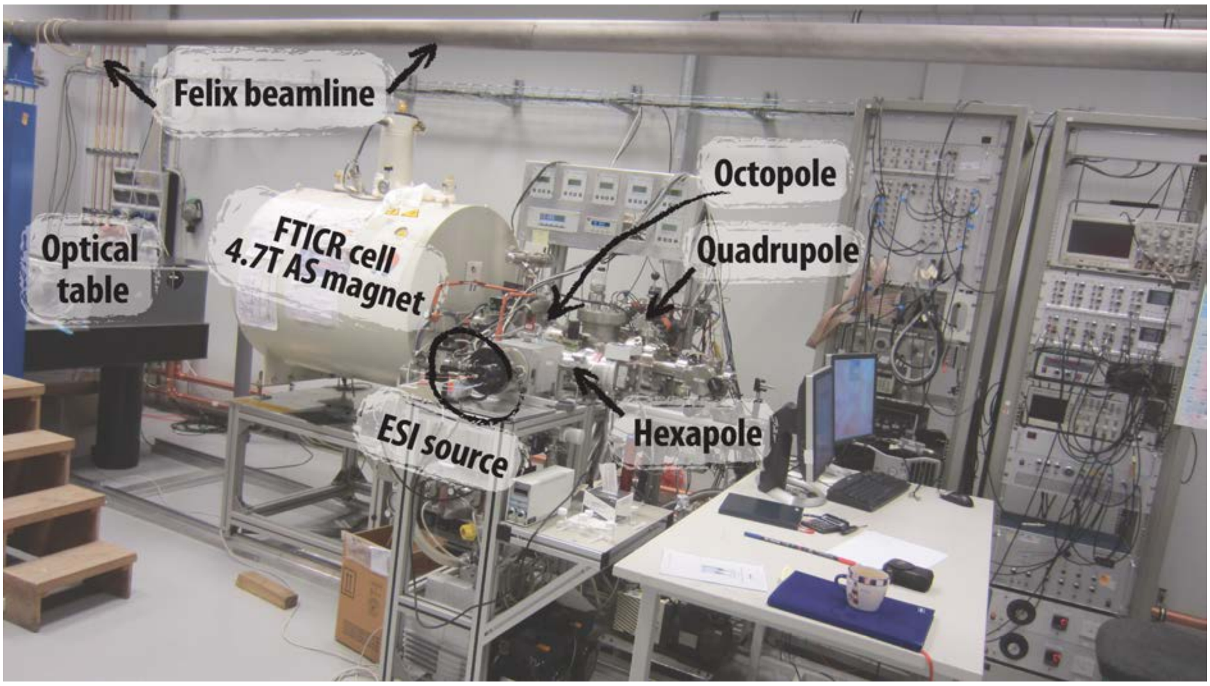


Figure 2.15. FTICR lab at the FELIX facility Nijmegen. Overview of the experimental setup. From the FELIX beamline, light can be transported to our optical table using a gold flip mirror. The FELIX beam can be focused such that the photon flux and the overlap with the trapped ions inside the ICR cell gives the highest (IRMPD) fragmentation yield. Water vapor in air absorbs some of the IR radiation which decreases the number of photons that are reaching the ICR cell. This is avoided by "flushing" the optical table box with nitrogen.

drawn into the mass spectrometer by the potential difference between the needle and the cone (analyzer entrance) as depicted in Figure 2.9. In this source, the ion trajectory is bent in a z-shape following the lines of the electric field while neutrals, which do not feel the electric field, and thus are unaffected, flow on toward the first-stage vacuum pump. The advantage of a z-spray source compared to the conventional "linear" electrospray sources is that there is a limitation for solvent molecules and other neutrals to enter the mass spectrometer or to build up on the cone, avoiding excessive contamination and thus enhancing the sensitivity and performance of the instrument.

The vacuum system of the FTICR-MS consists of a chain of differentially pumped vacuum chambers, leading from the atmospheric pressure ESI ion source to the high vacuum in the ICR cell (10^{-8} – 10^{-9} Torr). The main chamber is evacuated with a 500 l/s turbo molecular pump (Pfeiffer Vacuum Co.). The ICR high vacuum is separated from the higher pressure regions in the inlet of the apparatus with a gate valve. The vacuum system of the instrument presents a housing for the ion optics as depicted in Figure 2.17. With the ESI source in operation and the gate valve open, a pressure of around 10^{-6} Torr is maintained in the bender chamber by another 500 l/s turbo molecular pump. In the inlet section of the vacuum system, the lens chamber is at a pressure of around 10^{-4}-10^{-5} Torr (estimated), the hexapole chamber at 10^{-3} Torr and the nozzle-skimmer region of the ESI source at 1-4 Torr. These pressures are achieved by two 230 l/s turbo molecular pumps (Pfeiffer Vacuum Co.) mounted on the hexapole and ion lens chambers. The turbo pumps are backed by a set of rotary vane pumps of 20 m^3/hr.

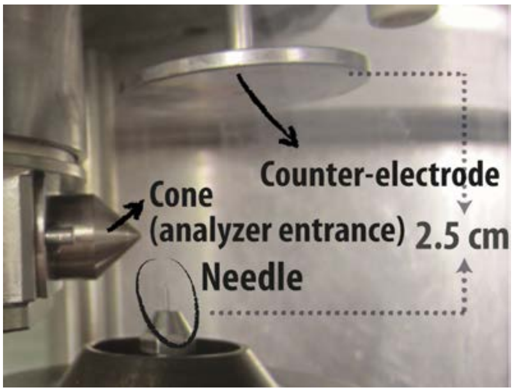


Figure 2.16. Z-spray source- electrospray chamber. The advantage of a z-spray source compared to conventional "linear" electrospray is that the amount of solvent molecules and other neutrals entering the mass spectrometer is limited, so that the ion source does not need to be cleaned frequently and the sensitivity and performance of the instrument remain constant for longer periods of time.

Ion optics. Figure 2.14 presents a sketch of the ion transfer from the ion source to the ICR cell. After the ions are formed in the ESI source, ions are guided by RF and electrostatic ion optics. The ion journey in this particular setup involves accumulation in a linear hexapole trap (Figure 2.18), focusing and steering by an ion lens and deflection plates, a 90-degree turn using a quadrupole deflector (bender) and guiding using a 1-m long RF octopole until they reach the ICR cell.

The 24 cm long RF hexapole trap is used to accumulate ions before pulse injection into the FTICR cell. The ions are axially trapped by the DC potentials applied to the skimmer and on the end cap located upstream and downstream from the hexapole, respectively. The DC bias voltage in combination with the RF amplitude then determine the effective potential well depth of the trap. Since the pressure in this chamber is around $3 \cdot 10^{-3}$ Torr, the ions undergo many collisions during accumulation and the well depth determines the collision energy. The hexapole accumulation trap can thus be used as a collision cell, where fragment ions are formed from the precursor ion. The collision induced dissociation (CID) process will

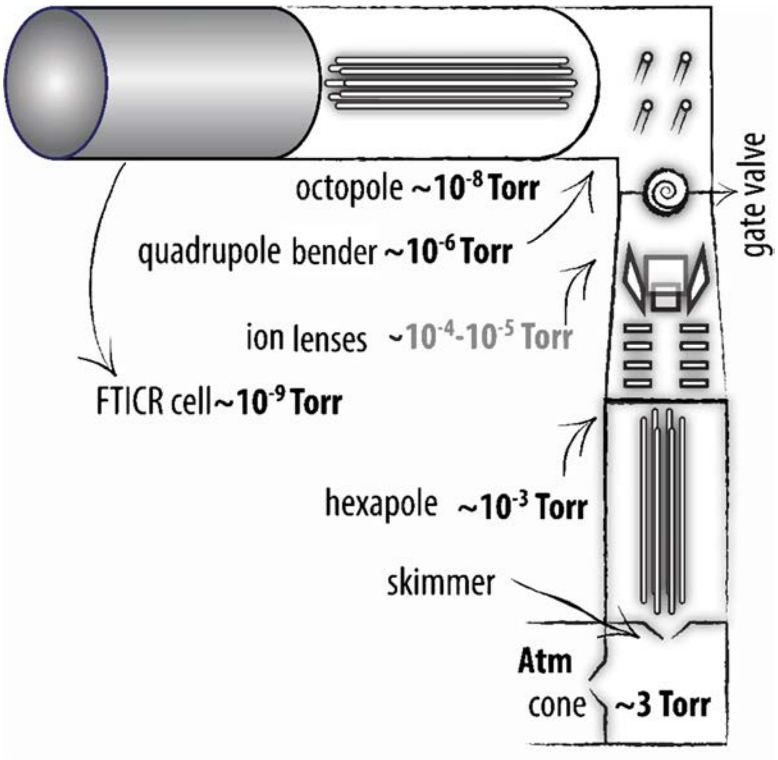


Figure 2.17. The vacuum system of the FTICR-MS consists of a series of differentially pumped vacuum chambers and presents housing for the ion optics. From atmosphere, the pressure drops to about 2-3 Torr in nozzle-skimmer region of the ESI source, dropping further in the hexapole chamber to 10^{-3} Torr. In the lens chamber it is around 10^{-4}-10^{-5} Torr (estimated), while in the bender region the pressure goes to approximately 10^{-6} Torr. After the bender, ions reach the high vacuum region of $10^{-8} – 10^{-9}$ Torr. When the instrument is not operating, the ion lens and bender regions are separated from each other with the gate valve.

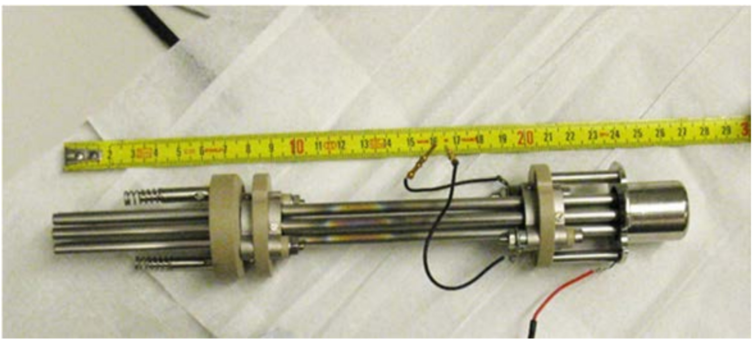


Figure 2.18. Hexapole The hexapole is used to accumulate ESI-generated ions prior to their injection into the ICR cell. The end cap electrode (right) has a 2 mm hole for ion transmission. Depending on values of the DC bias, trapping voltages and RF amplitude, the hexapole can also be used to induce collisional dissociation.

be described in more detail in the next chapter. The hexapole operates at an RF frequency of approximately 1.6 MHz with amplitudes up to 260 V peak-to–peak. The DC bias is set to a value around 35 V (polarity changes with the ions` charge) and determines the kinetic energy of the ions travelling downstream to the ICR cell.

After the hexapole, an Einzel lens is used to focus the ions. Two sets of deflection plates are used to steer the ions in the up/down and left/right directions. Subsequently, ions are deflected by 90 degrees using a quadrupole bender (Extrel GP-111), bringing them on-axis with the magnetic field. To transfer the ions from the deflector to the ICR cell, a RF octopole ion guide is used. Use of the octopole as a final ion guide stage has the advantage to circumvent the magnetic mirror effect (the magnetic mirror effect is a tendency of a charged particles to bounce back from the high field region)[51]. Moreover, we use the octopole ion guide to "switch" the DC reference voltage of the ions, reducing their kinetic energy from the initial 35 eV, before they arrive at the ICR cell. This method is illustrated in Figure 2.19. The ions are extracted from the accumulation stage at a bias voltage of about 35 V into the 1 m long RF octopole, which is initially biased at 0 V, thus giving the ions a kinetic energy of 35 eV as they enter the guide (Figure 2.19, A). Once the ions are inside the octopole, the DC bias voltage is switched to about -30 V (Figure 2.19, B). Since this bias voltage is on the full length of octopole rods, there is no change in the DC electric field along its axis, so that the ions do not gain or lose kinetic energy from this bias voltage switch; however, their potential energy is modified. Once the ions exit the octopole ion guide, they "see" the ICR cell, which is biased at ground, and by climbing the potential difference they slow down to a kinetic energy of about 5 eV. The ions can now conveniently be trapped by applying a moderate voltage on the ICR trapping plates, switched at the appropriate moment to capture the ions.

A simple electronic circuit, triggered by a TTL (transistor–transistor logic) pulse from the FTICR-MS data acquisition program, was constructed to apply the DC potential to the octopole rods having an RF frequency of about 1 MHz. Using this potential switch, re-referencing of the kinetic energy of the ions avoids the use of a pulsed buffer gas for capturing the ions in the ICR cell. This also significantly improves the duty cycle of the experiment as no pump delays are introduced. In addition, no collisional heating occurs, which has important advantages for the spectroscopic investigation of ions; lower internal energy results in a higher spectral resolution.

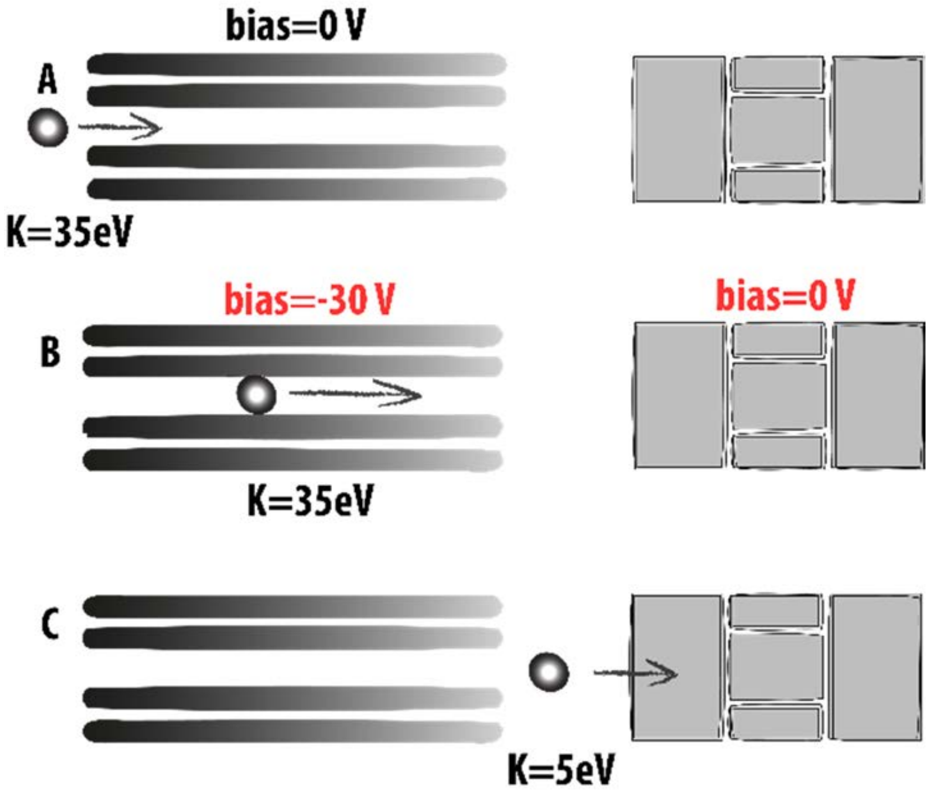


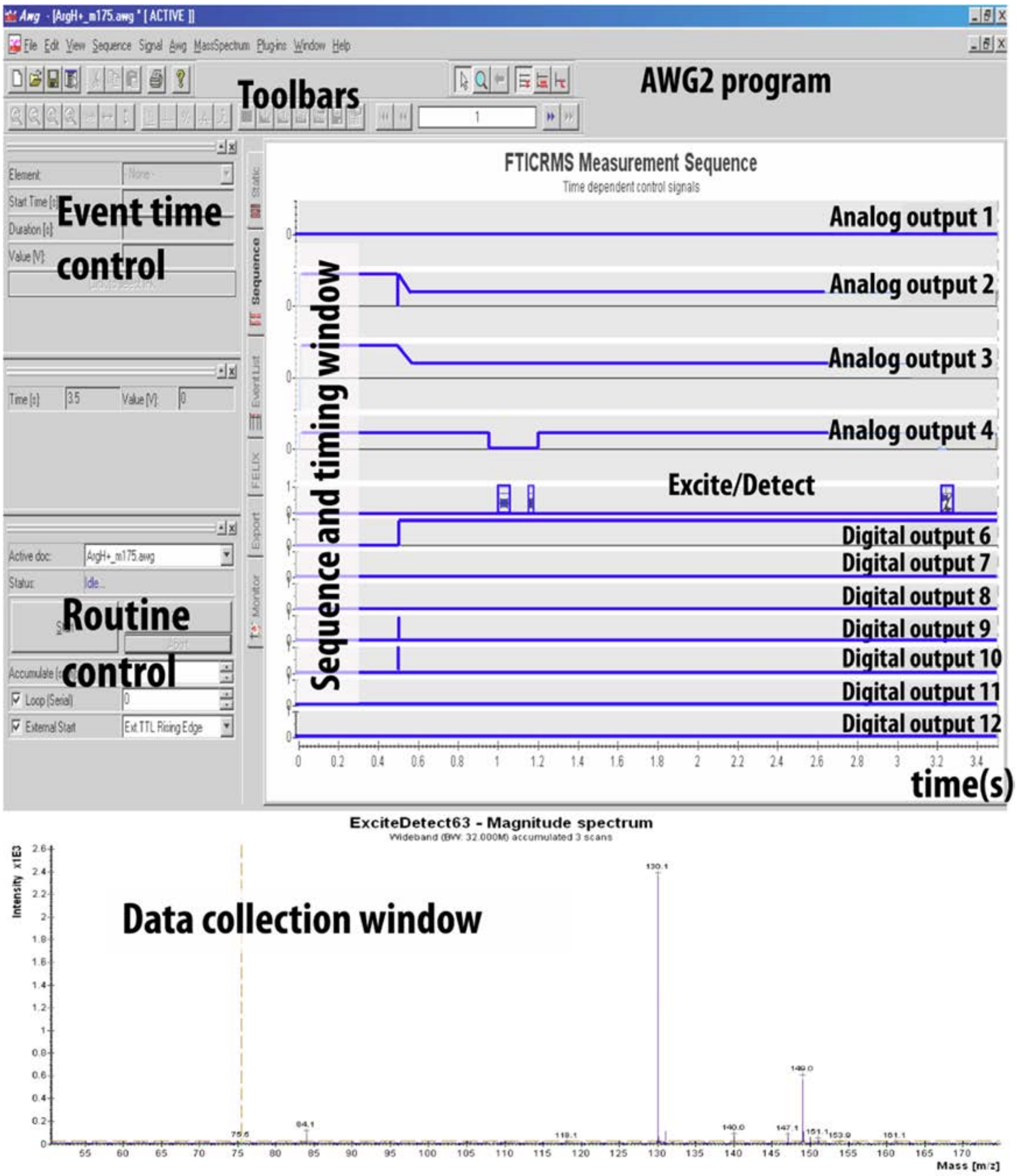
Figure 2.19. Schematic sequence of events used to guide and trap externally generated ions in the ICR. (A) A bunch of ions extracted from the accumulation hexapole with a kinetic energy of typically 35 eV is sent into a 1-m long RF-driven octopole ion guide. (B) Once inside the octopole, the DC bias on the octopole rods is abruptly switched to –30 V, thereby lowering the potential energy of ions. However, the ions' kinetic energy does not change since there is no DC potential gradient in the forward direction at any moment. (C) Upon leaving the octopole, the ground potential on the ICR cell (right) reduces the ions' kinetic energy to typically 5 eV so that they can be trapped easily. Note that the values given are typical and are actually optimized experimentally. Adapted from ref [51].

Once in the ICR cell, the ions can be mass isolated by a stored waveform inverse Fourier transform (SWIFT) excitation pulse[52,53], ejecting unwanted ions from the ICR cell. This method uses frequency specific RF excitation to accelerate undesired ions to a radius that is larger than that of the ICR trap. The ions thus hit the walls and are neutralized. The mass selectivity of FTICR MS is necessary for spectroscopy as well as for tandem mass spectrometry experiments. Both of these methods rely on recording fragment ion signals to generate the final experimental data. If there would be more than one ionic species present inside the ICR cell at the beginning of the experiment, it may not be possible to predict which of the fragments formed upon IRMPD belong to which of the precursor ions.

For the detection of a broad mass range, a frequency-sweep broadband excitation (RF chirp) is used. An arbitrary waveform generator (AWG) is programmed to sweep over frequencies corresponding to the molecular weights in the range of interest, typically from a few tens of kHz to MHz over a few tens of ms (typically 0.02 s). This causes all ions to expand their radius of cyclotron motion and allows their detection as explained in Section 2.1.1 of this chapter.

The FELIX (or OPO) infrared beam is introduced into the vacuum chamber of the mass spectrometer via a KRS-5 (thallium bromoiodide) IR transparent window (see Figure 2.14). The IR laser beam is aligned to pass parallel to the ICR cell, reaching a copper mirror (CM1) from where it is reflected to a second copper mirror, through the ICR cell. The inside of the copper excitation electrodes of the ICR cell are polished so that they act as a multipass reflection cell for the IR beam. This setup increases the overlap with the ion cloud in the ICR cell and improves the dissociation efficiency by a factor of approximately three, relative to a single-pass geometry. A second window (barium fluoride, BaF_2), mounted on the same flange but on-axis, allows the use of additional light sources in the experiment (e.g. UV/VIS or CO_2 laser). Both of the windows can be exchanged depending on the needs of an experiment. However, use of the optical beam path that is aligned over the copper mirrors restricts the use of some lasers (e.g. UV).

Control of each of the components in the FTICR is largely done manually by tuning the electronics (e.g. DC voltages, RF generators, amplifiers, etc.). The aim of tuning the instrument is to increase the ion signal intensity, and it is normally repeated for each experiment. The tuning can be divided into four sections related to the ESI source, hexapole ion accumulation, ion optics and event timings. Tuning the ESI source includes varying the needle, cone and skimmer voltage as well as nitrogen and sample solution flow. The high voltage needle supply (FUG Elektronik GmbH, Rosenheim, Germany, HCP 14-6500) can provide HV up to 6.5 kV, however for the majority of ions, values between 1.8 and 4 kV achieve the best signal intensities. The cone and skimmer voltages are on the order of tens to 200 V. Here it is important to mention the significance of the actual potential difference between the cone and the skimmer; for ions to be transferred from cone to skimmer, the skimmer needs to be on a lower potential. Secondly, by using a higher potential difference (e.g. above 30-40V) between the cone and skimmer, the kinetic energy of an ion can be increased, efficiently inducing fragmentation in this short but high-pressure (around 3 Torr) section of the ion path. Nitrogen and sample flow rates can be adjusted to optimize signal intensity. Important settings for the hexapole include the RF amplitude, DC bias, and the voltage on the end cap electrode. For ions to enter the hexapole, its bias must be lower than that at the skimmer. A typical value of the hexapole bias is around 35 V. The hexapole exit voltage is the actual potential barrier that causes ions to accumulate in hexapole. Triggered by the event timing, this potential drops to a value lower than the trap bias at the desired point in the experimental sequence. Tuning the ion optics can improve the transmission of ions from the source into the ICR cell. Typically these adjustments are made at the beginning of the day and if the source polarity is changed. Adjustments of the ion lenses between samples are minor. When changing the polarity of the instrument (switching from positive to negative ion mode or vice versa), the polarity of the needle HV, skimmer, cone, ion lenses, bender, hexapole bias and trapping electrodes needs to be reversed.



Time [s] 3.5 Value [V] 0

Active doc: ArgH+_m175.awg

Status: Idle...

☑ Loop (Serial) 0

☑ External Start Ext TTL Rising Edge

Figure 2.20. Basic layout of the AWG program interface. The routine control panel is used to control selection, stopping and starting an experimental routine as well as changing the number of accumulation scans. The event time control panel is used to change timing and duration of events in the routine. The sequence and timing window is used for setting and monitoring different channels (analog and digital outputs) over time. For example, by observing the excite/detect channel we can see that there are two excite pulses at 1 s and 1.2 s (for isolation) and that excitation/detection takes place at 3.2 s of this particular sequence. The mass spectrum of each cycle (or average) is presented in the data collection window.

Control Software (AWG). The data acquisition and experiment control is managed via a software package adapted from Heeren and co-workers.[54] The control system consists of a PC graphical user interface (GUI) for design of pulse sequences and initiation of the experiment (illustrated in Figure 2.20). This system consists of a PC-based platform for measurement and automation – a PXI (National instruments PXI-1042 Series) rack loaded with DAC (digital-to-analog converter) and TTL boards, providing analog and digital control pulses with microsecond timing, an arbitrary waveform generator (AWG), for ICR excitation and a digitizer to record the time-domain transient generated by the coherent ion motion. The sequence, including all non-static signals, is shown in a graphical timeline as presented in Figure 2.20. Via this GUI all timings and some voltage values are set. A simplified sequence of the experiment is presented in Figure 2.21. Using the AWG program, the timing of the hexapole and octopole triggers (digital outputs 9 and 10 respectively), the time dependent voltage on the trapping electrodes (analog output HQ 1 and 2), the SWIFT and excite/detect pulses (AWG/TD), the time of irradiation from FELIX or the OPO, (Digital output 11) are controlled. Digital outputs 8 and 12 can be used for additional pulses such as adding a CO_2 laser pulse or use of SORI-CID (sustained of resonance irradiation CID, which uses an RF pulse on the excitation electrodes for in-cell ion fragmentation; since it was not used for the experiments in this thesis, it will not be explained in more detail). The length of the sequence and number of averages depends on the experimental requirements, but the length is typically from 3 to 15 seconds, with three (or more) mass spectra averaged for each IR wavelength.

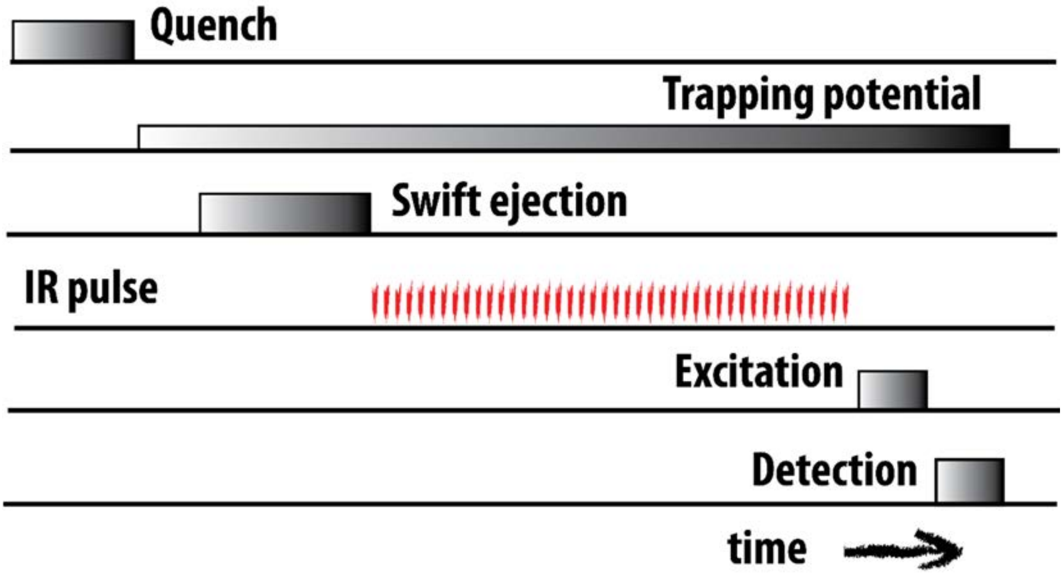


Figure 2.21. Simplified time sequence of events in the FTICR-FELIX experiments. Quench is used to empty the ICR cell of all the ions from the previous sequence. Another set of ions is trapped inside the ICR cell and swift ejection is used for mass isolation. A Trapping potential is applied during the entire sequence except while quenching. This is because the same trapping electrodes are used for both trapping and quenching (see explanation 2.1.1. I). Mass isolated ions can be irradiated with IR photons. During the IR multiple photon dissociation (IRMPD) process, fragments are formed. All ions formed inside the cell are then excited using RF, and subsequently detected as the recorded interferogram. The quenching, swift, excitation and detection events are on the order of milliseconds. The IR irradiation is on the order of seconds. (All timings also depend on the specific experimental requirements)

The arbitrary waveform editor (AWE) program is used to inspect and post-process raw FTICR data and to design excitation waveforms for use in AWG (specifically for the construction of SWIFT pulses used for ion isolation). The AWE editor and an example of SWIFT generation are presented in Figure 2.22.

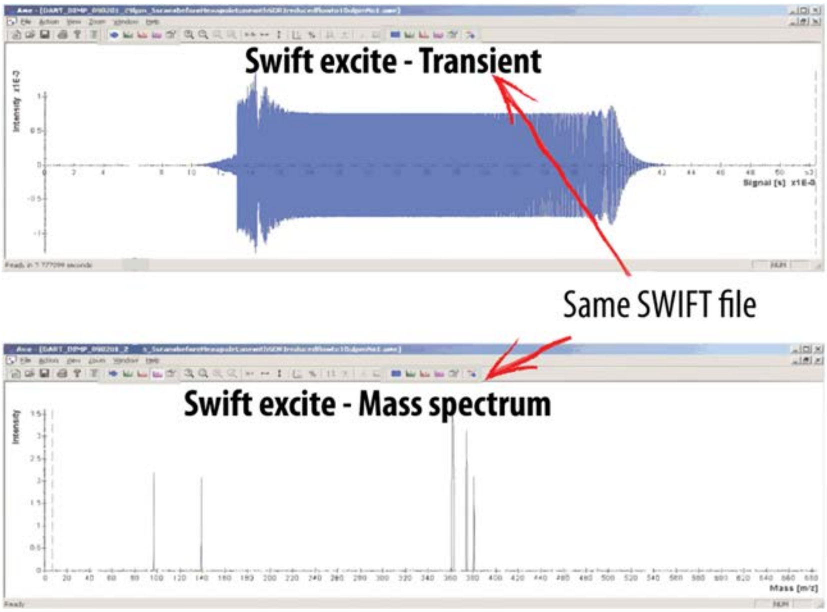
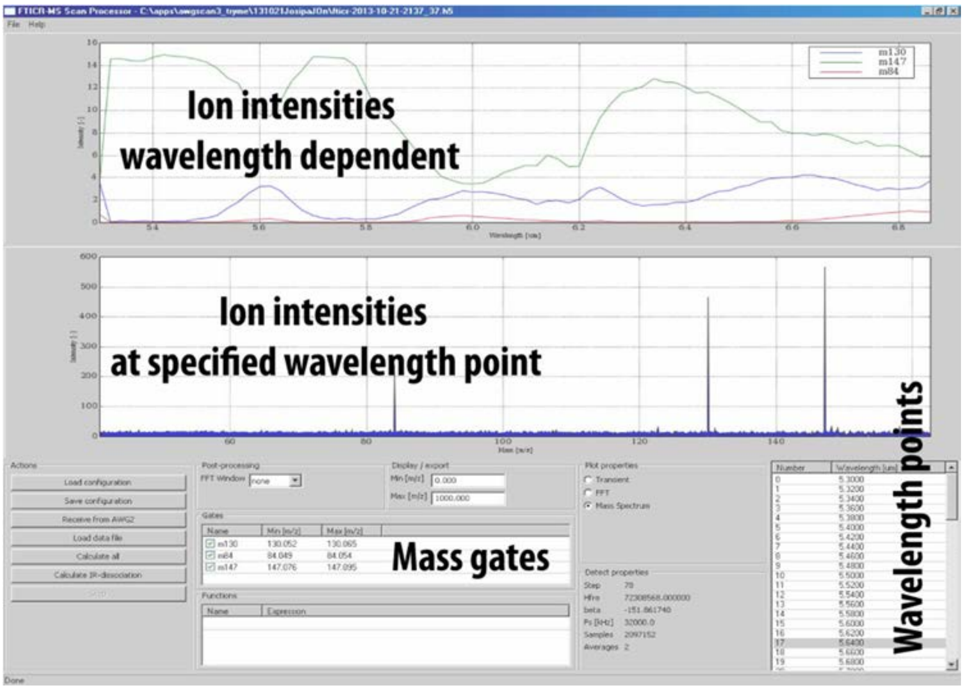


Figure 2.22. AWE editor program interface. This program allows selecting unwanted mass peaks for ejection, as presented in the bottom panel, and creates a SWIFT time-domain transient as presented in the upper panel. This SWIFT waveform is then loaded as an isolation event in the main sequence in the AWG program. At the appropriate time in the sequence, the software then sends the waveform to the amplifier connected to the excite plates of the ICR cell. SWIFT stands for stored waveform inverse Fourier transform.



Post-processing
FFT Window none

Display / export
Min [m/z] 0.000
Max [m/z] 1000.000

Gates

Name	Min [m/z]	Max [m/z]
☑ m130	130.052	130.065
☑ m84	84.049	84.054
☑ m147	147.076	147.095

Detect properties
Step 70
Mhz 72308568.000000
beta -151.861740
Fs [kHz] 32000.0
Samples 2097152
Averages 2

Number	Wavelength [um]
0	5.3000
1	5.3200
2	5.3400
3	5.3600
4	5.3800
5	5.4000
6	5.4200
7	5.4400
8	5.4600
9	5.4800
10	5.5000
11	5.5200
12	5.5400
13	5.5600
14	5.5800
15	5.6000
16	5.6200
17	5.6400
18	5.6600
19	5.6800

Figure 2.23. Overview of the FTICR-MS scan processor program interface. In the upper panel, specific ion intensities as selected by the mass gates are monitored over the time as a function of the wavelength. In the lower panel a mass spectrum for a chosen wavelength point can be displayed. This experiment shows the acquisition of IR spectra for protonated glutamine (mass 148) which dissociates into fragments at 84 and 147 m/z. The parent ion intensity drops when photon absorption followed by IRMPD occurs, and consequently, fragment ions are observed. The mass spectrum in the lower panel represents signal intensities of ions at a selected wavelength (5.64 μm). By plotting the fragmentation yield as a function of the wavelength one obtains the IR spectrum.

To generate IR spectra of mass selected ions, recording the signal intensity of the parent ion and/or the fragment ion(s) is necessary. The program used for this is FTICR-MS scan processor and its interface is presented in Figure 2.23. Briefly, this program allows setting the mass gates of monitored ions and records the intensity in each of these mass gates at each wavelength point. During an experiment, the program runs simultaneously with AWG so that the IRMPD spectrum that is being recorded is visualized "on the fly". After the wavelength scan, the data is exported and processed using commercial data analysis software such as Origin, where power and wavelength calibration of the recorded spectrum is performed.

2.3.2. FELIX

FELIX (Free Electron Laser for Infrared eXperiments) is a user facility and versatile source of radiation in the infrared and far infrared regions. A schematic overview of FELIX as it was constructed at the FOM institute Differ (Rijnhuizen) is presented in Figure 2.24.[38] Currently FELIX operates at Radboud University Nijmegen and an overview of the new FELIX bunker is presented in Figure 2.25. The boundary conditions set by a slightly new design and a "sister" FEL, FLARE, (FLARE is designed to lase from 100 to 1500 μm) make the FELIX overview and specifications slightly different from what they were at the FOM Institute (see the Figure 2.24 caption for modification details). Presently, FELIX operates in a wavelength range between 2.7 and 140 μm.

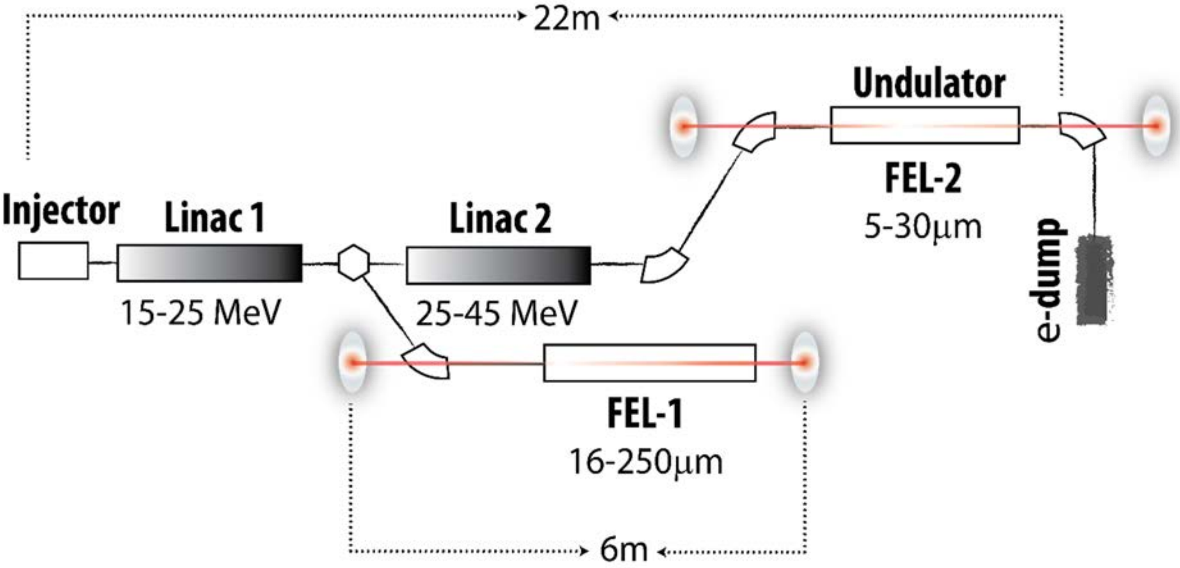


Figure 2.24. Layout of FELIX. The 3.8-MeV electron pulses are being accelerated using 3-GHz travelling-wave radio-frequency linear accelerators (RF LINACS) to energies of 15-25 and 25-45 MeV, respectively. Behind each LINAC, the electron beam can be bent and injected into the undulator where IR light is produced. This FELIX layout is from 1995 at the FOM Institute Differ (former Rijnhuizen). Since then, certain changes have been made; the capabilities of LINAC 2 have been improved and electrons can be accelerated to 50 MeV; FEL-2 is bent in a U-turn (see Figure 2.25); a new undulator is installed for FEL-1 (see Figure 2.25).

Figure 2.25. FELIX facility at Radboud University Nijmegen. The FELIX (*Free Electron Lasers for Infrared eXperiments*) facility in Nijmegen is an international user facility providing tunable radiation of high brightness in the mid- and far-infrared as well as the THz regime. The FELIX user facility features four beamlines for generation of infrared and THz radiation. The beamlines, FLARE, FELIX-1 and FELIX-2, provide continuously tunable radiation in the spectral range 2.7-1500 μm with peak powers up to 100 MW in picosecond pulses. The third beamline, FELICE, is designed for intra-cavity experiments on gas phase strongly bound systems. It provides continuously tunable radiation in the spectral range of 5-100 μm and the infrared intensity available to the experiment is up to one hundred times higher than available from FELIX.

The FELIX injector consists of an electron gun, prebuncher and buncher, producing 3.8 MeV electron pulses. These electrons are further accelerated by one or two RF linear accelerators (LINACs) to energies of 15-25 MeV or 25-45 MeV, respectively. After either LINAC, the electron beam can be bent into the FEL resonator. Both FELIX 1 and FELIX 2 have undulators consisting of two rows of samarium-cobalt permanent magnets forming 38 field periods of 65 mm length. The tunability of FELIX is achieved by varying the K value of the undulator by changing the distance between the two arrays of magnets (i.e. changing the magnetic field strength). Two gold mirrors at each end of the undulator constitute the optical resonator. After the undulator, the electrons are bent and finally dumped. A fraction of the light produced is coupled out through a small hole in one of the mirrors and transported through a beam transport system to one of the experimental user stations. The optical beam line is evacuated to eliminate absorptions by atmospheric gases, in particular water vapour and carbon dioxide, and allows for a maximum beam diameter of 70 mm. Infrared transparent windows separate the machine vacuum from the optical beam line.

Due to the use of RF linear accelerators, FELIX is a pulsed laser. The electron beam consists of a train of micropulses and thus, the optical output is also composed of a series of micropulses. The approximately picosecond micropulses are spaced by 1 or 40 ns (Figure 2.26), reflecting the frequency of the RF. The operational requirements of the LINACS limit the train of micropulses to approximately 10 μs, forming a macropulse. The repetition rate of FELIX macropulses is normally 5 or 10 Hz. The FELIX emits intense radiation with pulse energies up to 100 mJ per macropulse. The peak power during a micropulse is on the order of megawatts. For the experiments that will be presented in this thesis, the tunability and high power of FELIX were of crucial importance.

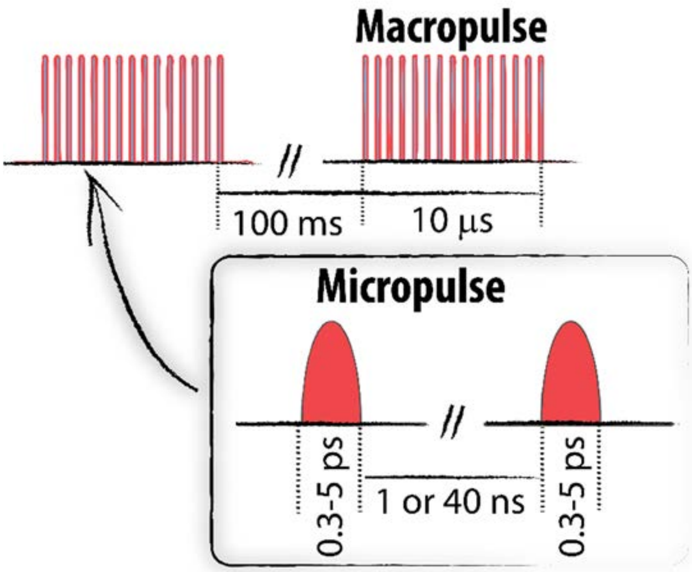


Figure 2.26. The FELIX light pulse shape. Length and repetition rate of a macropulse is a direct function of the LINAC. For LINAC resonators made of metal, heating by the RF field means that the pulses must come in a short burst (a few microseconds), followed by a long period of cooling (10's of milliseconds). The micropulse length is a function of the desired wavelength i.e. the number of optical cycles.

As a unique light source, FELIX has found various experimental applications and the list of research projects is continuously expanding; this includes time-resolved spectroscopy, non-linear optics, spectroscopy of molecules and ions in the gas phase, charge carrier dynamics in condensed matter physics, biophysics, quantum information and many others.

2.3.3. IR-OPO (at the FELIX facility)

The OPO/OPA (LaserVision, Bellevue, WA, USA) at the FELIX facility is designed to convert the fixed frequency radiation of a Nd:YAG pump laser source (Innolas Spitlight 600, Munich, Germany) into a tunable output beam in the mid-IR range with a bandwidth of approximately 3 cm^{-1}.

A schematic overview of the pulsed OPO laser is presented in Figure 2.27. The pump laser provides the OPO with 4-6 ns, 600 mJ pulses at 1064 nm. The pump beam is split into two beams, one going to the OPO and one going via a delay line to the OPA. The beam going to the OPO is frequency-doubled to 532 nm using a potassium titanyl phosphate crystal (KTP) (second harmonic frequency generator stage). The half-wave plate controls the efficiency of the conversion so that the 532 nm beam intensity can be varied. After frequency-doubling of the 1064 nm beam, the residual 1064 nm radiation is dumped. The 532 nm beam enters the OPO cavity where, using KTP crystal, signal ($v_{s,1}$) and idler ($v_{i,1}$) beams are produced. The idler from the OPO is combined with the delayed 1064 nm pump beam in the OPA stage. Difference-frequency mixing in the four counter-rotating potassium titanyl phosphate arsenate (KTA) crystals amplifies the signal (idler from the OPO stage) and

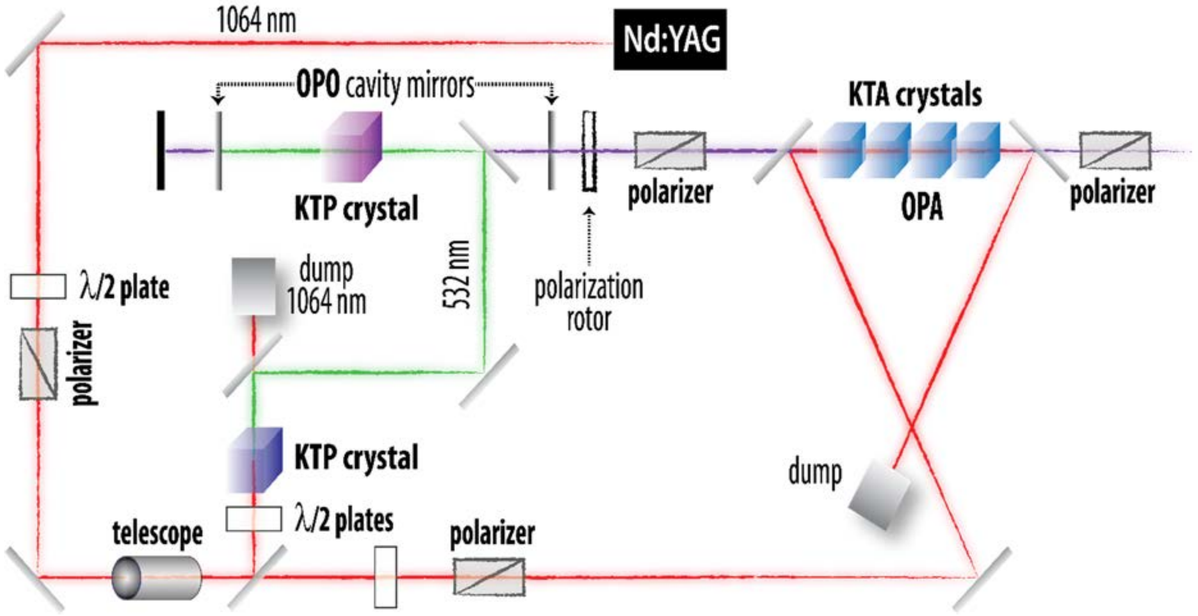


Figure 2.27. OPO at the FELIX facility. Tunable light generation from this particular OPO/OPA system is achieved by doubling the frequency of an Nd-YAG laser (1064 nm to 532 nm) using a KTP crystal, which is followed by the OPO stage where photons of lower frequency are generated according to equation 2.26. The idler from the OPO is combined with the 1064 nm pump (delayed) beam and difference-frequency mixed in the OPA (composed of four potassium titanyl phosphate arsenate (KTA) crystals). After the OPA, the remaining 1064 nm pump radiation is removed with a dichroic filter and a separate Brewster 'stack of plates' polarizer is used to isolate either the idler, or the signal. A set of polarizers along the beam line is used to control the light polarization, since for their mixing, the beams` polarization relative to each other and relative to the crystal axis is of crucial importance. A telescope is used to match the beam profile to the crystal size.

generates the idler in the 3 µm range. After the OPA, the remaining 1064 nm pump radiation is removed with a dichroic filter and a separate Brewster 'stack of plates' polarizer is used to isolate either the idler, $v_{i,2}$, or the signal, $v_{s,2}$, which have vertical and horizontal polarization, respectively.

The frequencies of the produced light beams can be calculated using this set of the equations:

Signal and idler - OPO stage $$v_{i,1} + v_{s,1} = v_{532} \tag{2.26}$$

Signal and idler - OPA stage $$v_{i,2} + v_{s,2} = v_{1064} \tag{2.27}$$

and, since $v_{i,1} = v_{s,2}$, the light that leaves this OPO/OPA system will have frequency:

$$\begin{aligned} v_{i,2} = v_{1064} - v_{i,1} = \ v_{1064} - (v_{532} - v_{s,1}) = v_{s,1} - v_{1064} \\ = \ v_{s,1} - 9394.8\ cm^{-1} \end{aligned} \tag{2.28}$$

This system is wavelength calibrated using a wavemeter measuring the wavelength of $v_{s,1}$ and v_{1064}.

The phase-matching of the beams is achieved by angle-tuning of the crystals. The parametric system is designed in such a way that all crystals are rotated in the plane parallel to the optical bench for safety reasons. Crystal positions in the OPO and OPA are controlled remotely by a computer so that the OPA output intensity can be optimized. Each motor controlling a crystal angle can be individually calibrated to a tenth–order polynomial expansion of position as a function of a wavelength. Pulse energies of the described OPO are between 5 and 20 mJ at 10 Hz repetition rate and pulse duration of 5 ns, as defined by pump source. The wavelength range used for the experiments of this thesis is from 2.6 to 3.5 µm.

REFERENCES

(1) Thompson, J. J. *Rays of positive electricity* Proceedings of the Royal Society **1913**, *A 89*, 1.

(2) Lawrence, E. O.; Livingston, M. S. *The production of high speed light ions without the use of high voltages* Phys. Rev. **1932**, *40*, 19.

(3) Ernst, R. R.; Anderson, W. A. *Application of Fourier Transform Spectroscopy to Magnetic Resonance* Rev. Sci. Instrum. **1966**, *37*, 93.

(4) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Fourier transform ion cyclotron resonance mass spectrometry: a primer* Mass Spectrom. Rev. **1998**, *17*, 1.

(5) Marshall, A. G.; Comisarow, M. B.; Parisod, G. *Relaxation and spectral line shape in Fourier transform ion resonance spectroscopy* J. Chem. Phys. **1979**, *71*, 4434.

(6) He, F.; Hendrickson, C. L.; Marshall, A. G. *Baseline Mass Resolution of Peptide Isobars: A Record for Molecular Mass Resolution* Anal. Chem. **2000**, *73*, 647.

(7) König, M.; Bollen, G.; Kluge, H. J.; Otto, T.; Szerypo, J. *Quadrupole excitation of stored ion motion at the true cyclotron frequency* Int. J. Mass Spectrom. Ion Proc. **1995**, *142*, 95.

(8) Ledford, E. B.; Rempel, D. L.; Gross, M. L. *Space charge effects in Fourier transform mass spectrometry. II. Mass calibration* Anal. Chem. **1984**, *56*, 2744.

(9) McIver, R. T.; Hunter, R. L.; Bowers, W. D. *Coupling a quadrupole mass spectrometer and a Fourier transform mass spectrometer* Int. J. Mass Spectrom. Ion Proc. **1985**, *64*, 67.

(10) Kofel, P.; Allemann, M.; Kellerhals, H.; Wanczek, K. P. *External generation of ions in ICR spectrometry* Int. J. Mass Spectrom. Ion Proc. **1985**, *65*, 97.

(11) Comisarow, M. B. *Comprehensive Theory for Ion Cyclotron Resonance Power Absorption: Application to Line Shapes for Reactive and Nonreactive Ions* J. Chem. Phys. **1971**, *55*, 205.

(12) Dunbar, R. C. *Transient Ion Cyclotron Resonance Method for Studying Ion–Molecule Collision and Charge Transfer Rates:* N_2^+ *and* CH_4^+ J. Chem. Phys. **1971**, *54*, 711.

(13) Amster, I. J. *Fourier Transform Mass Spectrometry* J. Mass Spectrom. **1996**, *31*, 1325.

(14) Senko, M. W.; Hendrickson, C. L.; Paša-Tolić, L.; Marto, J. A.; White, F. M.; Guan, S.; Marshall, A. G. *Electrospray Ionization Fourier Transform Ion Cyclotron Resonance at 9.4 T* Rapid Commun. Mass Spectrom. **1996**, *10*, 1824.

(15) Schaub, T. M.; Hendrickson, C. L.; Horning, S.; Quinn, J. P.; Senko, M. W.; Marshall, A. G. *High-Performance Mass Spectrometry: Fourier Transform Ion Cyclotron Resonance at 14.5 Tesla* Anal. Chem. **2008**, *80*, 3985.

(16) Morris, H. R.; Panico, M.; Barber, M.; Bordoli, R. S.; Sedgwick, R. D.; Tyler, A. *Fast atom bombardment: A new mass spectrometric method for peptide sequence analysis* Biochem. Biophys. Res. Commun. **1981**, *101*, 623.

(17) Yamashita, M.; Fenn, J. B. *Electrospray ion source. Another variation on the free-jet theme* J. Phys. Chem. **1984**, *88*, 4451.

(18) Karas, M.; Hillenkamp, F. *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons* Anal. Chem. **1988**, *60*, 2299.

(19) Zeleny, J. *Instability of Electrified Liquid Surfaces* Phys. Rev. **1917**, *10*, 1.

(20) Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *Molecular Beams of Macroions* J. Chem. Phys. **1968**, *49*, 2240.

(21) Fenn, J. B.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Electrospray ionization for mass spectrometry of large biomolecules* Science **1989**, *246*, 64.

(22) Yamashita, M.; Fenn, J. B. *Negative ion production with the electrospray ion source* J. Phys. Chem. **1984**, *88*, 4671.

(23) Loeb, L. B.; Kip, A. F.; Hudson, G. G.; Bennett, W. H. *Pulses in Negative Point-to-Plane Corona* Phys. Rev. **1941**, *60*, 714.

(24) Kebarle, P.; Verkerk, U. H. *Electrospray: From ions in solution to ions in the gas phase, what we know now* Mass Spectrom. Rev. **2009**, *28*, 898.

(25) Taylor, G. I.; McEwan, A. D. *The stability of a horizontal fluid interface in a vertical electric field* J. Fluid Mech., *22*, 1.

(26) Fernandez de la Mora, J. *The Fluid Dynamics of Taylor Cones* Annu. Rev. Fluid Mech. **2007**, *39*, 217.

(27) Rayleigh, L. *XX. On the equilibrium of liquid conducting masses charged with electricity* Philosophical Magazine Series 5 **1882**, *14*, 184.
(28) Cole, R. B. *Some tenets pertaining to electrospray ionization mass spectrometry* J. Mass Spectrom. **2000**, *35*, 763.
(29) Kebarle, P. *A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry* J. Mass Spectrom. **2000**, *35*, 804.
(30) Fernandez de la Mora, J. *Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism* Anal. Chim. Acta **2000**, *406*, 93.
(31) Iribarne, J. V.; Thomson, B. A. *On the evaporation of small ions from charged droplets* J. Chem. Phys. **1976**, *64*, 2287.
(32) Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L. *Combined Charged Residue-Field Emission Model of Macromolecular Electrospray Ionization* Anal. Chem. **2008**, *81*, 369.
(33) Curley, R. *The Britannica Guide to Inventions That Changed the Modern World*; Rosen Publishing Group, Incorporated, 2009.
(34) Maiman, T. H. *Stimulated optical radiation in ruby* Nature **1960**, *187*, 493.
(35) Hirshfield, J.; Bernstein, I.; Wachtel, J. M. *Cyclotron resonance interaction of microwaves with energetic electrons* Quantum Electronics, IEEE Journal of **1965**, *1*, 237.
(36) Pantell, R. H.; Soncini, G.; Puthoff, H. E. *Stimulated photon-electron scattering* Quantum Electronics, IEEE Journal of **1968**, *4*, 905.
(37) Madey, J. M. J. *Stimulated Emission of Bremsstrahlung in a Periodic Magnetic Field* J. Appl. Phys. **1971**, *42*, 1906.
(38) Oepts, D.; van der Meer, A. F. G.; van Amersfoort, P. W. *The infrared free-electron laser facility FELIX* Infrared Phys. Technol. **1995**, *36*, 297.
(39) Zhirong, H.; Ingolf, L. *Free-electron lasers: SACLA hard-X-ray compact FEL* Nature Photonics **2012**, *6*, 505.
(40) Mcneil, B. W. J.; Thompson, N. R. *X-ray free-electron lasers* Nature **2010**, *4*, 814.
(41) Bonifacio, R.; De Salvo, L.; Pierini, P.; Piovella, N.; Pellegrini, C. *Spectrum, temporal structure, and fluctuations in a high-gain free-electron laser starting from noise* Phys. Rev. Lett. **1994**, *73*, 70.
(42) Oomens, J.; Tielens, A. G. G. M.; Sartakov, B. G.; Helden, G. v.; Meijer, G. *Laboratory Infrared Spectroscopy of Cationic Polycyclic Aromatic Hydrocarbon Molecules* Am. Astron. Soc. **2003**, *591*, 968.
(43) Galué, H. A.; Rice, C. A.; Steill, J. D.; Oomens, J. *Infrared spectroscopy of ionized corannulene in the gas phase* J. Chem. Phys. **2011**, *134*.
(44) Oomens, J.; Sartakov, B. G.; Tielens, A. G. G. M.; Meijer, G.; Helden, G. v. *Gas-Phase Infrared Spectrum of the Coronene Cation* Astrophysical J. **2001**, *560*, L99.
(45) Jaeqx, S.; Oomens, J.; Cimas, A.; Gaigeot, M.-P.; Rijs, A. M. *Gas-Phase Peptide Structures Unraveled by Far-IR Spectroscopy: Combining IR-UV Ion-Dip Experiments with Born–Oppenheimer Molecular Dynamics Simulations* Angew. Chem. Int. Ed. **2014**, *53*, 3663.
(46) Pagel, K.; Kupser, P.; Bierau, F.; Polfer, N. C.; Steill, J. D.; Oomens, J.; Meijer, G.; Koksch, B.; von Helden, G. *Gas-phase IR spectra of intact a-helical coiled coil protein complexes* Int. J. Mass Spectrom. **2009**, *283*, 161.
(47) Almasian, M.; Grzetic, J.; Berden, G.; Bakker, B.; Buma, W. J.; Oomens, J. *Gas-phase infrared spectrum of the anionic GFP-chromophore* Int. J. Mass Spectrom. **2012**, *330–332*, 118.
(48) Prell, J. S.; Chang, T. M.; Biles, J. A.; Berden, G.; Oomens, J.; Williams, E. R. *Isomer Population Analysis of Gaseous Ions From Infrared Multiple Photon Dissociation Kinetics* J. Phys. Chem. A **2011**, *115*, 2745.
(49) Ross, A. J.; Dreiocker, F.; Schäfer, M.; Oomens, J.; Meijer, A. J. H. M.; Pickup, B. T.; Jackson, R. F. W. *Evidence for the Role of Tetramethylethylenediamine in Aqueous Negishi Cross-Coupling: Synthesis of Nonproteinogenic Phenylalanine Derivatives on Water* J. Org. Chem **2011**, *76*, 1727.
(50) Woodin, R. L.; Bomse, D. S.; Beauchamp, J. L. *Multiphoton dissociation of molecules with low power continuous wave infrared laser radiation* J. Am. Chem. Soc. **1978**, *100*, 3248.
(51) Polfer, N. C.; Oomens, J.; Moore, D. T.; von Helden, G.; Meijer, G.; Dunbar, R. C. *Infrared Spectroscopy of Phenylalanine Ag(I) and Zn(II) Complexes in the Gas Phase* J. Am. Chem. Soc. **2005**, *128*, 517.
(52) Marshall, A. G.; Wang, T. C. L.; Ricca, T. L. *Tailored excitation for Fourier transform ion cyclotron mass spectrometry* J. Am. Chem. Soc. **1985**, *107*, 7893.

(53) Guan, S.; Marshall, A. G. *Stored waveform inverse Fourier transform (SWIFT) ion excitation in trapped-ion mass spectometry: Theory and applications* Int. J. Mass Spectrom. Ion Proc. **1996**, *157–158*, 5.
(54) Mize, T. H.; Taban, I.; Duursma, M.; Seynen, M.; Konijnenburg, M.; Vijftigschild, A.; Doornik, C. V.; Rooij, G. V.; Heeren, R. M. A. *A modular data and control system to improve sensitivity, selectivity, speed of analysis, ease of use, and transient duration in an external source FTICR-MS* Int. J. Mass Spectrom. **2004**, *235*, 243.


wireless
LOVE
Sicko

Chapter 3.

Experimental and theoretical methods

In the previous chapter, details of the experimental tools used in this thesis were described. However, several questions remain: why are we able to record IR spectra of mass selected ions using FELIX and how can we "convince" an ion to fragment? Historically, IR spectroscopic structure-determination used known diagnostic vibrational frequencies of various functional groups to infer general structural information of analytes or unknown compounds. With emerging developments in analytical methods, this approach of determining molecular structures has largely been replaced with more modern methods, often involving theoretical chemistry. In this chapter, some of the physical principles behind these methods, including infrared multiple photon dissociation (IRMPD) spectroscopy, collisional induced dissociation (CID) and the various theoretical approaches, will be discussed.

3.1. Infrared spectroscopy – fundamentals

Even if the Romans were the first to notice the ability of the prism to disperse light into a rainbow, the basis for spectroscopy was systematically established with Isaac Newton's developments in optics in the mid-17th century. Newton applied the word "spectrum" to the rainbow of colors produced when white light is passed through a prism. Infrared radiation was discovered by Sir Wilhelm Friedrich Herschel in 1800.[1] Herschel directed sunlight through a glass prism and directed the different spectral regions onto a thermometer. Since he measured a response not only for the visible light, but also noticed a temperature increase for the region just beyond the red end of the visible spectrum, Herschel concluded that there must be an invisible form of light in this region. This was the form of light that we know today as infrared radiation.

In 1814, the first spectroscope was invented by Joseph von Fraunhofer. This development brought the experimental advances that enabled spectroscopy to become a more precise and quantitative scientific technique. Since that time, spectroscopy has played a significant role in chemistry, physics, astronomy and a variety of other scientific fields. One of the earliest founders of infrared spectroscopy was William W. Coblentz. When Coblentz entered Cornell University, infrared spectroscopy was in, what nowadays would be considered a rather primitive state. However, by 1905 he had obtained hundreds of infrared spectra (in the 1-15 µm range) by “point-by-point” measurements with a prism instrument of his own construction.[2]

Infrared spectroscopy is a technique that measures the interaction of matter with infrared radiation, determining the frequency of vibrations of the atoms within a molecule**.** As different molecules consist of a variety of functional groups, they exhibit different infrared activity in specific regions of the spectrum. Thus, IR spectroscopy is most commonly used for the identification of functional groups by measuring their associated molecular vibrations. In many cases, measuring the IR spectrum of a molecule allows one to propose a molecular structure. An infrared spectrum is normally obtained by passing infrared radiation through a sample and determining what fraction of the incident radiation is absorbed at each particular frequency. The energy at which a peak in an absorption spectrum appears corresponds to the frequency of a particular vibration in the sample molecule. The infrared region between approximately 600-2000 cm^{-1} is commonly referred to as the fingerprint region as many small organic molecules exhibit a unique and diagnostic pattern in this region.

Not all vibrational modes in a molecule are necessarily IR active. Only if there is a net change of the dipole moment (the dipole moment is a measure of the charge separation in a system) within a molecule during the vibrational motion and the frequency of the dipole oscillations is equal to the frequency of the incident radiation an IR photon absorption can occur. If there is no change in the dipole moment during the molecular vibration, that vibration is infrared "silent" or inactive. An example of an IR inactive vibration is the symmetric C=O stretching vibration of CO_2, since stretching and contracting of the two CO bonds is in a counteracting manner and does not result in breaking of symmetry and thus change of dipole moment (presented in Figure 3.1). The opposite is true for the asymmetric stretch of CO_2. When the C=O bonds are in equilibrium the dipole moment equals zero, but when one bond elongates and other one contracts the charge distribution is changing and results in a "vibrationally induced dipole".

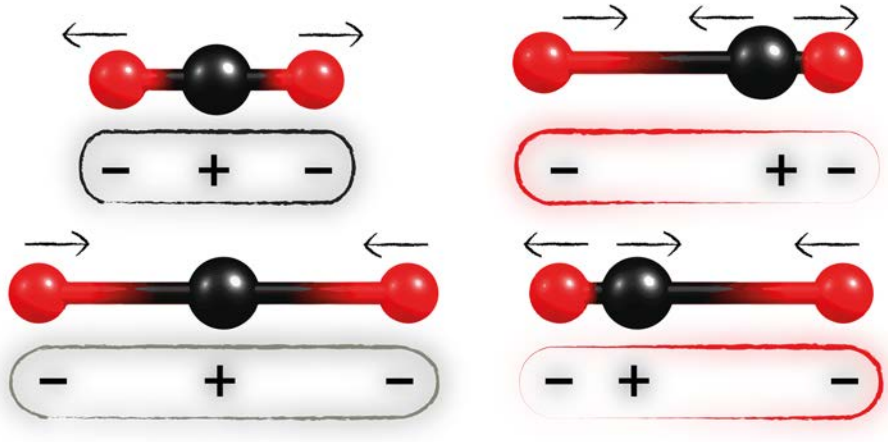


Figure 3.1. Vibrationally induced dipole moment. A criterion for IR absorption is a net change in dipole moment in a molecule as it vibrates. Considering a molecule such as CO_2, the charge distribution between carbon and oxygen is not evenly allocated since oxygen is more electronegative than carbon and has a higher electron density. The C=O bond thus has a large dipole moment and is thus polar. The dipole moment is determined by the magnitude of the charge difference and the distance between the two centers of charge. For the symmetric stretch (left) the dipole moment is not changing, thus this mode is IR "silent". During the asymmetric vibrational motion (right) the symmetry of the charge distribution within a molecule is changing inducing the dipole moment, thus this vibration is IR active.

An additional requirement for photon absorption is that the eigenfrequency of the vibration must be equal to the frequency of the incident radiation. This means that the energy difference between vibrational levels (E_1 and E_2) must be equal to the photon energy.

$$\Delta E = E_1 - E_2 = hv \tag{3.1}$$

The number of normal vibrational modes in a molecule can be calculated using simple equations based on the degrees of freedom in the molecule. For a linear molecule, we can attribute three degrees of translational freedom and two of rotational freedom. Additionally, atoms can move relative to each other along the bond axis and in and out of plane. These are defined as stretching and bending vibrational modes, respectively (Figure 3.2). In the case of a diatomic molecule, only stretching of the bond is possible which then accounts as one degree of vibrational freedom. All molecules have 3N degrees of freedom, where N is the number of atoms. For non-linear molecules, all rotational motions can be described in terms of rotations around three axes and there are three degrees of translational motion. This leaves 3N-6 degrees of freedom to constitute the vibrational motions. For a linear molecule however, rotation around the bond axis makes no spatial change and reduces the rotational degrees to two, leaving 3N-5 degrees of freedom for

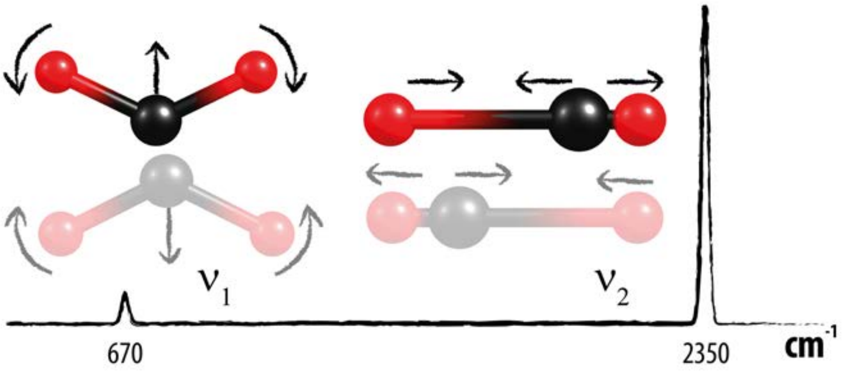


Figure 3.2. Vibrational modes of CO_2. The vibrational degrees of freedom for linear molecules can be calculated as 3N-5. Considering CO_2, four vibrations are expected. However, the IR spectrum features only two bands. This can be explained by the two bending modes at 670 cm^{-1} being degenerate (i.e. have the same frequency) and the symmetric stretching mode not resulting in a change of dipole moment. The difference in intensity is a direct result of the magnitude of the dipole change which is higher for the asymmetric stretching mode.

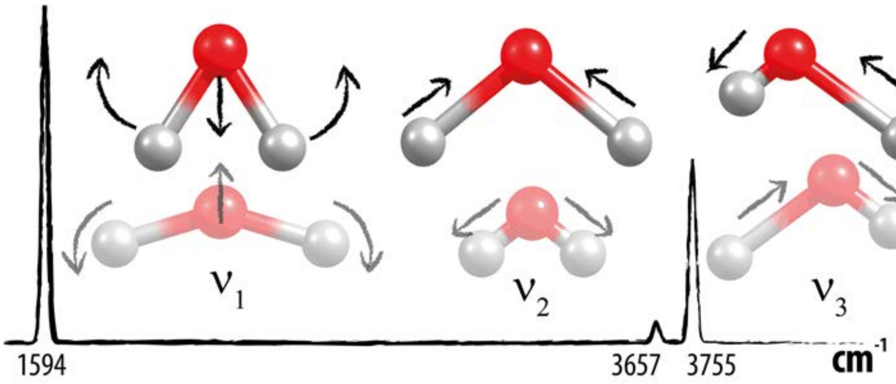


Figure 3.3. Vibrational modes of H_2O. The vibrational degrees of freedom for non-linear molecules can be calculated as 3N-6. In the case of H_2O, three vibrations are expected. In contrast to the CO_2 symmetric stretching mode, the symmetric stretching mode in the water molecule is infrared visible, since this vibration involves a dipole moment fluctuation.

vibrations. Examples of vibrational modes of linear and nonlinear molecules are depicted in Figure 3.2 and 3.3, respectively. Consequently, there are many different vibrations for even a relatively small molecule. As well, some complexity in an IR spectrum can arise from coupling of individual vibrations over either a part of or over the whole molecule.

Measured IR absorption bands are not infinitely narrow and there are several reasons for the broadening of these bands. Prominent for IR spectroscopy of gas phase systems is the Doppler effect. The Doppler effect refers to the change in the frequency of light emitted/absorbed from a body in motion. An additional reason results from the Heisenberg uncertainty principle which relates the uncertainties in the lifetime of an excited vibrational state and its energy:

$$\delta E = \hbar/\tau \tag{3.2}$$

Here, δE is the uncertainty in energy, ħ is the reduced Planck constant, and τ is the lifetime of an excited state. All excited states are subject to lifetime broadening and, the shorter the lifetimes of the states involved in a transition, the broader the corresponding spectral lines. In the absence of collisional de-excitation (i.e. only radiative decay), this is referred to as natural broadening. The natural linewidths of IR absorptions are very narrow, since in a collisionless environment the lifetime of a vibrationally excited state is typically very long (i.e. on the order of milliseconds or more). However, molecular collisions reduce the lifetime so that the bandwidth is affected (known as pressure broadening). For high vacuum regimes, time in between collisions is typically on the order of seconds and broadening due to collisions is, at least for the experiments presented in this thesis, negligible.

In our experiments, the most prominent contributions to the observed bandwidth arise from the envelope over the unresolved rotational sub-structure and broadening due to the ion heating upon multiple photon excitation. With the increase of the internal energy in the ion, the density of states increases and vibrational transition frequencies progressively experience anharmonic broadening and red shifting (see section 3.1.3). Additionally, a thermal distribution of ion conformers may exist, so that a superposition of similar IR spectra is observed resulting in further spectral congestion. However, in the observed IRMPD spectra, bands are usually still resolved well enough for structural assignment, confirming the presence or absence of structural moieties and functional groups within the molecule. For more detailed conformational analysis based on IR spectroscopy, different approaches, such as spectroscopy of cold molecular ions, have also been applied by various research groups[3].

The intensity of an absorption depends on the magnitude of the change of the dipole moment associated with the observed molecular vibration. The larger this change is, the more intense the absorption band will be. In mathematical terms, the absorption intensity is proportional to $[\partial M/\partial x]^2$, the square of the change in dipole moment, M, with respect to the change in the normal coordinate, x.[4] The frequency of these vibrations depends on the inter-atomic binding energy, which determines the force needed to stretch or compress a bond, and will be explained in following part of this chapter.

3.1.1. Molecular vibration - classical approach

Classically, the vibrational motion within a molecule can be represented using the harmonic oscillator model if we assume that two atoms in the molecule are connected with bonds acting as oscillating ideal springs (Figure 3.4). Within this approach, the intramolecular bonds are considered to be elastic, e.g. all atoms undergo small harmonic oscillations around their equilibrium position obeying Hooke's law. According to this law, $\vec{F}$ is the restoring force exerted by the spring when it's either elongated or compressed from its equilibrium position and can be written as:

$$\vec{F} = -k\vec{x} \tag{3.3}$$

Here, k is the force (spring) constant and $\vec{x}$ is displacement. The force exerted by the spring on each of the masses always points in the direction of the equilibrium position. This system exhibits oscillatory behavior. The masses move in opposite directions, and if they are not identical, the magnitude of displacement of m_1 and m_2 are different, such that the center of mass is stationary during the entire oscillation. The simplest approach to analyze the evolution of this motion is to use coordinates based on observing the system from the point that does not move, called the center of mass coordinate.

At this point, the physical picture of the system changes from two masses connected by a

$$x = \frac{m_1x_1 + m_2x_2}{m_1 + m_2} \tag{3.4}$$

spring, to a single mass, called the reduced mass:

$$\mu = \frac{m_1m_2}{m_1 + m_2} \tag{3.5}$$

According to Newton`s second law, balance of forces leads to

$$\vec{F} = -k\vec{x} = m\vec{a} = m\frac{d^2x}{dt^2} \tag{3.6}$$

From here, we can describe the motion of the oscillator as $x(t) = A\sin(\omega t + \varphi)$. When we solve Equation 3.6 with reduced mass instead of m as written previously, we derive that the angular frequency of the motion is

$$\omega = \sqrt{k/\mu} = 2\pi/T \tag{3.7}$$

The classical expression for the potential energy of a harmonic oscillator, for a given force, can be derived as the integral of the force over the displacement.

$$V(x) = -\int F dx = {}^1\!/_2\, kx^2 \tag{3.8}$$

From here we conclude that the frequency of the IR absorption band depends on the masses of atoms involved and the strength of the chemical bond, as characterized by the force constant, k. The higher the masses of the atoms in the system are, the lower is its vibrational frequency. This relation is useful in the assignment of vibrations involving hydrogen atom motions; replacing the hydrogen atom(s) by deuterium, a red shift of the vibrational frequency by a factor of $\sqrt{2}$ is expected. This is referred to as the isotope effect. The constant k is highly dependent on the strength of a chemical bond. This is exemplified by comparison of the approximate positions of the vibrational stretches for single, double and triple carbon-carbon bonds at approximately 1100, 1650 and 2150 cm^{-1}, respectively. Of course, the complete electronic structure of a molecule, as well as its environment plays

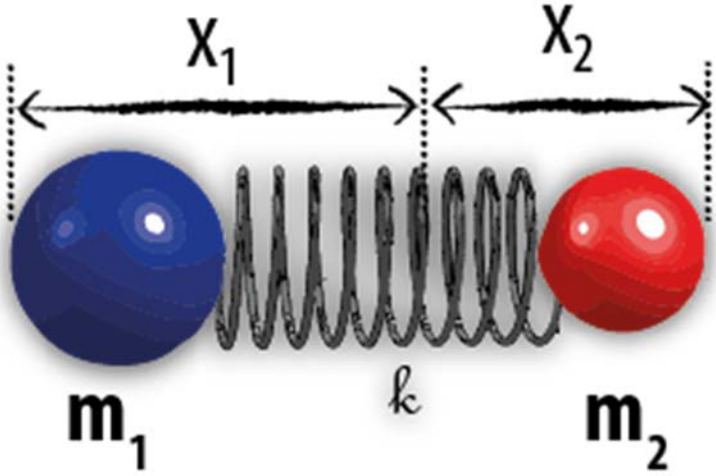


Figure 3.4. Harmonic oscillator model. If the bond joining two atoms of masses m_1 and m_2 acts as a perfect spring, the two masses move back and forth as the bond compresses and expands with a characteristic frequency. The force constant (k) reflects how strongly the bond resists being stretched or compressed.

an important role in the bond strengths and consequently in the precise vibrational frequencies.

3.1.2. Molecular vibration – quantum mechanical approach

The Schrödinger equation for a 1-dimensional harmonic oscillator may be obtained by using the classical spring potential from Equation 3.8. The time independent Schrödinger equation then becomes

$$\frac{-\hbar^2}{2\mu}\frac{d^2\Psi(x)}{dx^2}+\frac{1}{2}kx^2\Psi(x)=E\Psi(x) \tag{3.9}$$

To first approximation, the vibrational motions of a polyatomic molecule can be described as harmonic oscillators. In this harmonic approximation, it can be shown that the 3N – 6 vibrational degrees of freedom are strictly independent 1-D harmonic oscillators. Hence, the Schrödinger equation can be solved for each normal coordinate independently, giving the energy levels as

$$E_i=h\nu_i\left(n_i+\frac{1}{2}\right)=\frac{h}{2\pi}\sqrt{\frac{k_i}{\mu_i}}\left(n_i+\frac{1}{2}\right) \tag{3.10}$$

Here, n_i is the vibrational quantum number for the *i*-th normal coordinate, where $n_i = 0, 1, 2, \ldots$ Within this description, energy levels in a given mode are evenly separated by multiples of $h\nu_i$, and the zero-point energy, $E_{0,i} = h\nu_i/2$, is the energy of the ground vibrational state ($n_i = 0$). In the harmonic approximation, transitions are allowed between adjacent vibrational levels of a single vibrational mode, that is, the vibrational quantum number, n_i, can only change by one ($\Delta n_i = \pm 1$). Thus, the quantum-mechanical treatment shows that the vibrational energy of a molecule is not continuously variable but quantized. A parabolic potential, centered at the equilibrium bond distance (r_e), describes the potential energy (Figure 3.5).

Although the fundamental vibrational properties of molecules can often be understood within the harmonic model, it cannot account for all properties of real molecules, such as dissociation. The description of the potential energy by a quadratic function suggests an infinitely increasing restoring force as the bond length becomes larger, preventing the molecule from dissociation even for infinitely long bond distances. This assumption is therefore only adequate when the displacement from the equilibrium bond length is small (Figure 3.5). However, for large displacements (and especially at energies approaching the dissociation limit) this is not the case. At high internuclear distances, the restoring force levels off to zero (the molecule dissociates) and the potential reaches a plateau, which represents the dissociation energy.

Correcting for the deviation from the harmonic behavior of a molecular oscillator can be accomplished through the assumption of an anharmonic vibrational potential well. A popular anharmonic potential is the empirical Morse potential, which for a diatomic molecule is given by

$$V = D_e[1 - e^{-\sqrt{k/2D_e}(R-R_e)}]^2 \tag{3.11}$$

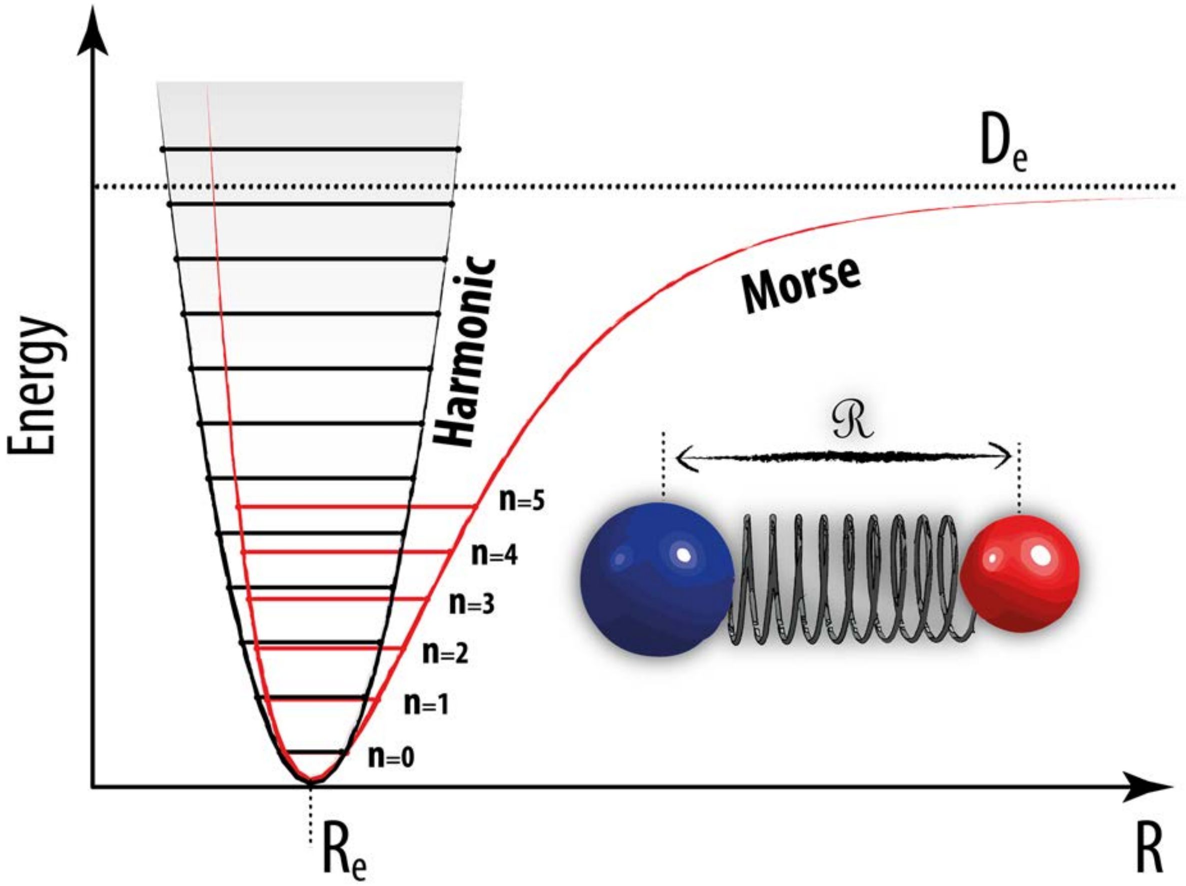


Figure 3.5. The Morse potential (red) and the harmonic oscillator potential (black). Unlike the energy levels of the harmonic oscillator potential, which are evenly spaced by $h\nu$, the Morse potential level spacing decreases as the energy approaches the dissociation energy. The Morse approximation is a better approach for representing the vibrational structure of the molecule than the harmonic oscillator model since it explicitly includes the effects of bond dissociation.

The Morse potential function is illustrated in Figure 3.5. At small internuclear distances, R, the Coulomb repulsion of the nuclei is high and dominates the description of the potential energy. At the equilibrium spacing, R_e, the potential energy between the atoms is most attractive and similar to the harmonic potential. However, when increasing the internuclear distance far beyond R_e, the restoring force $(-dV/dR)$ approaches zero. This leads to dissociation of the bond, where the dissociation energy D_e is the limit of the potential function at infinite separation of the nuclei. Solving the Schrödinger equation with this potential function can be shown to result in a vibrational level spacing which is no longer constant, but decreases with increasing vibrational quantum numbers (Figure 3.5). The corrected energy levels are given by equation 3.12.

$$E_i = h\nu_i\left(n_i + \frac{1}{2}\right)\left[1 - \chi(n_i + \frac{1}{2})\right] \qquad 3.12$$

The anharmonicity constant χ is a small positive number, so that the level spacing decreases as one climbs the vibrational ladder towards the dissociation limit.

An additional consequence of the anharmonicity is that the normal modes in a polyatomic molecule are no longer strictly independent (i.e. orthogonal). Rather, small couplings exist between them, which play a critical role in infrared multiple-photon dissociation chemistry as used throughout this thesis to probe molecular vibrations. Furthermore, mixing of normal vibrational modes can lead to the appearance of combination bands $(\nu_i \pm \nu_j)$ and overtones due to the $\Delta n = \pm 1$ selection rule being no longer strictly binding.

3.1.3. Infrared multiple photon dissociation (IRMPD)

Direct absorption infrared spectroscopy of ions in the gas phase is known to have a low sensitivity due to the inherently low number densities of ions of the same charge in the gas phase. A direct absorption experiment measures the attenuation of radiation transmitted through the sample. According to the Lambert-Beer law, the absorbance (A_ν) of the sample is given by

$$A_\nu = \ln\frac{I_\nu}{I_0} = \sigma_\nu n l \qquad 3.13$$

Here, σ_ν represents the total cross section, n the number of density of molecules and l the path length. In the FT-ICR apparatus, the maximum number density of ions of the same charge is limited by the space-charge limit, estimated around ~10^6-10^8 ions/cm^3. l is equal to the trap length (for example 20 cm) and σ_ν = ~10^{-18}-10^{-20} cm^2. Considering the above values, a change in the light intensity of ~10^{-10} is expected. This value is too low to be detectable. Therefore, other methods are necessary for the measurement of infrared spectra of gaseous ions, which are generally referred to as action spectroscopy methods. These methods measure effect on the molecule ("action") produced by their interaction with the radiation, rather than the attenuation of the transmitted light beam.

IRMPD spectroscopy is a form of action spectroscopy in which molecular fragmentation resulting from resonant photon absorption is monitored. Schematically it can be presented as in Expression 3.14.

$$AB^{+/-} \xrightarrow{nh\nu_{IR}} (AB^{+/-})^* \rightarrow A^{+/-} + B \ or \ A + B^{+/-} \qquad 3.14$$

Infrared multiple photon dissociation (IRMPD) spectroscopy provides a vibrational spectrum of gaseous ions through the observation of wavelength-dependent fragmentation of the irradiated molecules. Monitoring the extent of fragmentation induced by a tunable IR source allows the use of a mass spectrometer as a highly sensitive detector for vibrational spectroscopy measurements. In the mid-1970s, the groups of Dunbar, Beauchamp, Lee and Eyler were the first to use tunable IR lasers and the IRMPD method to measure vibrational spectra of gaseous ions.[5-10]

To dissociate typical covalent chemical bonds using IR photons, tens to hundreds of photons must be absorbed by a single ion to provide the minimum amount of energy required for dissociation (at a dissociation threshold of, for example, 3 eV, at least 24 photons at 10 µm would be required). However, due to the anharmonicty, the decreasing vibrational level spacing leads to a slow loss of resonance as multiple resonant photons are absorbed (Figure 3.6). Especially for systems with high dissociation energies (of a few eV) the coherent stepwise absorption fails to induce dissociation of molecular systems in the type of experiments presented in this thesis. Therefore, coherent multiphoton excitation, where all photons are absorbed in the vibrational ladder in a stepwise process ($n_i = 0 \rightarrow n_i = 1 \rightarrow n_i = 2\ldots$), is unrealistic at moderate laser intensities (often referred to as the anharmonicity bottleneck[11,12]). Thus anharmonicity appears to exclude IRMPD as an efficient excitation mechanism.

However, IRMPD can be accomplished by sequential incoherent absorption of photons. In the first stage of this mechanism, the state densities are low and energy levels are well separated; the effect of anharmonicity is also small. The IRMPD mechanism begins with the absorption of a resonant photon, increasing the internal energy of the molecule and rapidly leading to an increase in the density of vibrational-rotational states. The levels can no longer be considered as entirely discrete but rather form a quasi-continuum[13]. Although the quasi-continuum suggests that photons of any wavelength can be absorbed, this region is nevertheless characterized by semi-resonant absorption in zones near the original fundamental transition. This is followed by internal vibrational redistribution (IVR) which proceeds via the anharmonic coupling between normal modes. The vibrational energy of the excited resonant or “bright” state is redistributed into anharmonically coupled background or “dark” vibrational states. This allows for the system to absorb photons resonant with the fundamental transition of the bright state again, followed again by a redistribution of the energy. The internal energy is continually increased by each photon absorption until the dissociation threshold is reached and fragment ions are produced.

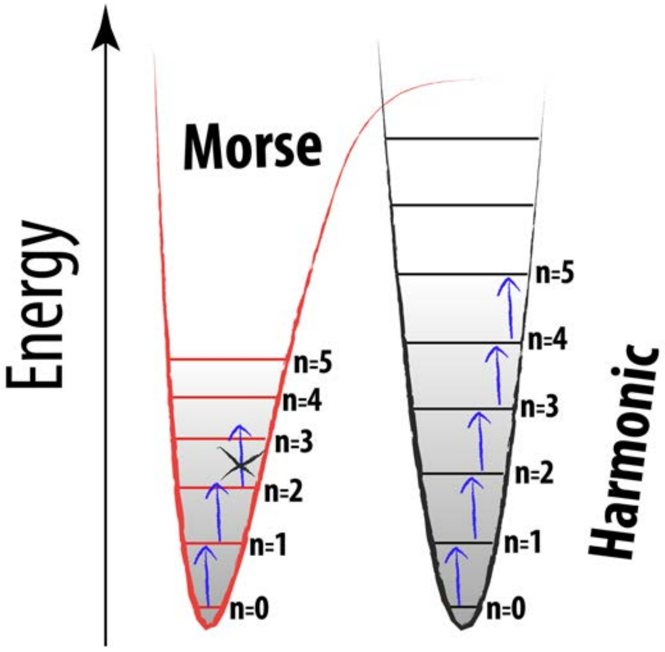


Figure 3.6."Ladder climbing" mechanism. To dissociate typical chemical bonds using IR photons, tens to hundreds of photons must be absorbed to provide the minimum amount of energy. In the harmonic approximation, this would be possible through "ladder climbing". However, for the more realistic Morse potential, vibrational anharmonicity hinders further resonant absorption of photons after a few transitions, anharmonic coupling between normal modes can explain IRMPD using the anharmonic approach.

The efficiency of IVR depends on the density of vibrational states ($\rho(E)$), and consequently, since it depends on the number of normal vibrational modes and vibrational energy, it is expected to be more efficient for larger systems. In small molecules, with a low density of vibrational states, the effect of the anharmonicity bottleneck is larger and higher laser intensities are necessary. Although a more sophisticated expression is necessary for an accurate description (especially for relatively low internal energies), the correlation between density of states and internal energy and size of the system is illustrated by Equation 3.15.

$$\rho(E) = \frac{1}{(s-1)!}\frac{E^{s-1}}{\prod_{i=1}^{s}\hbar\omega_i} \qquad 3.15$$

Where E represents the vibrational energy, s is the number of vibrational degrees of freedom and, ω_i, the frequency of the i^{th} vibrational mode.

When the density of states is large enough, IVR rapidly redistributes the internal energy into the bath of vibrational background states and allows absorption at the fundamental transition from $n_i = 0$ to $n_i=1$ again. The short lifetimes induced by IVR in combination with a small residual shifting of the $n_i = 1 \leftarrow 0$ transition with internal energy result in a broadening of absorption lines. For the majority of molecules, IVR lifetimes can be assumed on the order of picoseconds to one nanosecond[14,15], which corresponds to the micropulse spacing of FELIX. Thus, the excited vibrational mode is likely de-excited through dissipation of the energy into the bath of states before the next micropulse arrives. Since IVR allows high excitation of the system and induces short vibrational lifetimes, differences are observed with direct absorption spectra, especially slight red shifting of the IRMPD frequencies and possible broadening of the absorption bands. Nevertheless, well-resolved spectra are observed for the majority of small to medium-sized molecular ions.

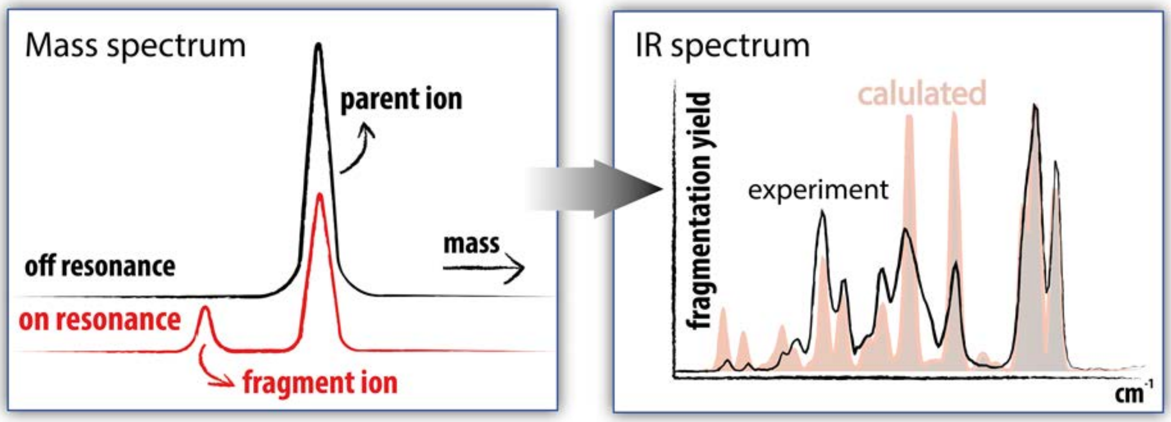


Figure 3.7. IRMPD experiment using MS detection. The left panel represents appearance of a mass spectrum of a selected ion when IR photons are off (black) and on (red) resonance. When photons are off resonance, there is no interaction, and only the parent ion signal is recorded. When IR radiation is on resonance with one of the vibrational modes within the selected ion, photon absorption occurs. The parent ion undergoes IRMPD which results in its signal intensity to decrease and that of the fragment ion to emerge (red). Recording the IRMPD efficiency (yield) as a function of IR wavelength results thus in an IR spectrum. The experimental result, an IR spectrum of the mass selected ion, is presented in the right panel (black). Still, this is just the first part of structural assignment. To determine the structure, theoretical spectra are calculated so that experiment can be compared to calculated spectra for several candidate structures.

Finally, a convenient way to present the degree of fragmentation of the parent ion to fragment ion(s) is the IRMPD efficiency (η_{IRMPD}) or yield

$$\eta_{IRMPD} = \frac{\sum I_{fragments}}{I_{parent} + \sum I_{fragments}} \qquad 3.16$$

Expressing the extent of fragmentation in this way suppresses fluctuations in the total ion current (TIC) signal from appearing in the final spectrum (for example, instabilities in the spray current from an electrospray source). A schematic presentation of recording IR spectra monitoring the IRMPD efficiency using a mass spectrometer is presented in Figure 3.7. It is also possible to record IRMPD spectra by monitoring only the depletion of a parent ion signal in a so-called “depletion” mode (in our experiments, for example, it is not so unusual to notice the electron detachment from anionic parent molecule as the only IRMPD channel). These spectra are inherently more sensitive to fluctuations in the parent ion signal and generally yield lower quality spectra. Often, this method requires several spectra to be recorded and averaged to obtain reliable spectral data.

Detailed structural analysis of IRMPD spectra is usually accomplished by a comparison to theoretical IR spectra of possible molecular geometries. Typically, this involves comparing IRMPD spectra to harmonic frequency calculations at the density-functional theory (DFT) level (Figure 3.7, right panel), which will be discussed later in this chapter.

3.1.4. Ion activation using multiple photon absorption

As described above, the IRMPD method with the use of a tunable laser can give us detailed structural information on the analysed ion. However, the aim of photo-dissociation experiments is not necessarily the recording of an infrared spectrum of an ion. Photo-dissociation can also be applied as an ion activation technique in MS/MS. With the developments of MS trapping instruments, such as quadrupole ion traps and the FTICR, it has become possible to store ions for long periods of time. Thus, ions can be irradiated, dissociated and the resulting fragments can be analysed. For experiments in this thesis, a continuous-wave CO_2 laser was occasionally used to generate fragment ions to be later spectroscopically analysed. In contrast to the irradiation of ions with a tunable laser exploiting resonant absorption, the efficiency of non-resonant absorption highly depends on the size of the system and thus for some ions use of this fragmentation method is impossible.

3.2. Collision induced dissociation (CID)

Within the last two decades, mass spectrometry has become the method of choice for protein identification. The crucial step in generating structural information of a protein or a peptide is analysing its dissociation products. Collision-induced dissociation (CID) is the most widely applied fragmentation method in tandem mass spectrometry (MS/MS) of peptides and proteins.

The first demonstration of gas-phase CID of molecular ions was obtained in some of the first mass spectra recorded by Sir J. J. Thompson with his mass spectrograph. Thompson noted in 1913 that positive rays (in particular H_2^+) in his ion deflection experiments grouped around certain values (which were later shown to correspond to H^- and H^+) on the experimental photograph taken[14]. Even if Thompson then did not actually claim that there is dissociation occurring caused by collisions, it is difficult to read his explanation in any other terms. Since that time, this phenomenon has been a subject of study throughout the development of MS during the twentieth century. This technique was introduced for the first time in 1968[15,16] for the detection, identification and structural assignment of molecules of any kind.

A crucial aspect of an MS/MS experiment is the reaction that occurs between the two MS stages. By far the most frequent reaction is unimolecular dissociation, normally achieved by some form of ion activation. Ion activation is necessary to raise the internal energy of the parent ion so that it dissociates before the analysis in the second MS stage. In CID, the precursor ion and the (inert) neutral collision partner, can be regarded as a projectile and target (Figure 3.8).

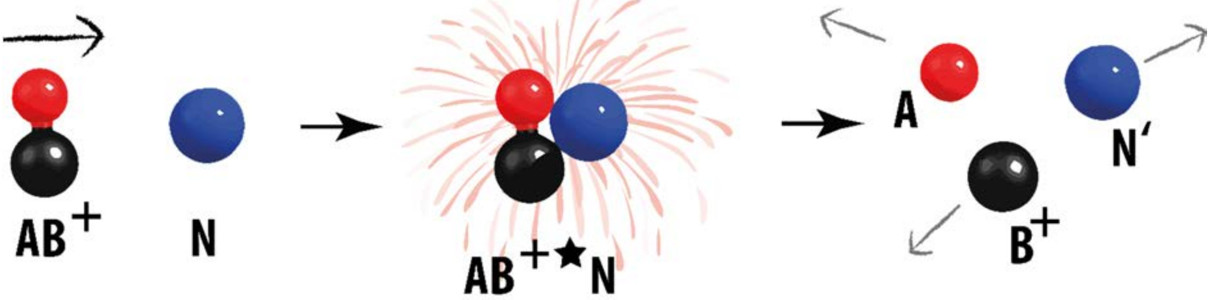


Figure 3.8. Collision induced dissociation (CID). In this cartoon representation of CID, the precursor ion and cell containing inert neutral, can be observed as a projectile and target. When an ion (AB^+) with high translational energy undergoes an inelastic collision with a neutral (N), part of the translational energy is converted into internal energy of the ion (AB^{+*}), leading to its fragmentation into neutral A and ion B^+.

The projectile ion dissociates as a result of a translational energy conversion to the ion's internal energy. This two-step mechanism can be expressed as

$$AB^+ + N \rightarrow [AB^+]^* + N' \qquad 3.17$$

$$[AB^+]^* \rightarrow A^+ + B \; or \; B^+ + A \qquad 3.18$$

Where $AB^+$is the precursor ion, N is the target neutral and A/B and $(A/B)^+$ are fragments.

There are several things to consider following one collision in the CID process and those are the magnitude of energy transfer, the distribution and variability of transferred energy, kinetics, and the efficiency.

The amount of energy that can be deposited into an ion depends on the relative collision energy of the two colliding bodies. Most CID experiments can be approximately classified into two energy regimes: high energy CID using keV kinetic energy, in the laboratory frame of reference, and low-energy CID with 1–100 eV kinetic energy in the laboratory frame. In peptide sequencing applications – and in the experiments in this thesis – low-energy collisions are typically used for fragment generation. The maximum energy that can be transferred to the precursor ion in the collision process is described by the following equation.

$$E_{com} = q \cdot V \frac{m_n}{m_n + m_i} \qquad 3.19$$

Here, q is the charge of the precursor ion, V is the acceleration voltage and m_n and m_i are the masses of the neutral and ion, respectively. In order to simplify the description, energy transfer in CID and the kinetic energy release during the dissociation of activated ions are deduced by applying classical conservation of momentum and energy. Therefore, it is convenient to work in the center-of-mass framework, where the center-of-mass momentum is zero, instead of the laboratory reference frame. In the latter, a collision is defined by the two separate particles involved, each having their individual position and velocity vectors. The product $q \cdot V$ represents the energy of an ion in the laboratory frame, E_{lab}, and E_{com} is its fraction, if the velocity of the neutral is ignored. From Equation 3.19 it is obvious that the efficiency of energy transfer is dependent on the mass of the ion and that the dissociation efficiency drops as the ion mass increases.

The number of collision events in the activation period also determines the overall amount of transferred energy. Most analytical applications of CID use collision regions operated at pressures high enough to ensure that numerous collisions occur, such that the internal energy is sufficiently increased to produce extensive dissociation. The number of collisions is determined by the collision frequency (ω_i) and the time period over which collisions may occur. The collision frequency depends on the collisional cross section, σ_i (if we approximate colliding bodies as hard spheres then $\sigma_i = \pi(r_a + r_b)^2$ where r_a and r_b are the radii; the cross-section is typically10^{-14}–10^{-17} cm^2), the relative velocity of the collision partners, $\vartheta_{i,G}$, and the gas pressure and temperature, P_G and T_G, respectively. Assuming ideal gas behavior, the collision frequency can be written as[17]

$$\omega_i = \sigma_i \vartheta_{i,G} \frac{P_G}{k_b T_G} \qquad 3.20$$

Increasing the pressure to obtain a larger collision energy transfer has its limitations, as it may compromise MS sensitivity. Additionally, as the number of collisions increases, the system can reach equilibrium such that further collisions are just as likely to remove energy from the ion as to add to it. In some cases, this collisional cooling at elevated pressures may actually increase the sensitivity and mass resolution[18,19].

The activation energy rapidly brings the system to higher vibrational (or electronic) states. Statistical theories of activated ion dissociation assume that internal energy equilibration, through all vibrational degrees of freedom, is faster than dissociation and experimental evidence shows that this holds for the vast majority of ions and molecules. Under multiple collision conditions, where energy is deposited in small increments at a variety of collision sites on the ion, collisional activation effectively gives rise to a statistical ion energy distribution. The fraction of the energy which is partitioned between the energy levels in the precursor ion can be written as

$$E_i = \frac{(E' + E_G + E_{com})D_i}{D_i + D_G + D_T} \qquad 3.21$$

where E' is the initial internal energy of the precursor ion, E_G is the internal energy of the precursor gas, D_i and D_G the number of degrees of freedom of the ion and the target molecule, respectively, and D_T represents the translational degrees of freedom[17]. For low collision energies, excitation is mostly vibrational, since the interaction time is on the order of 10^{-13}–10^{-14} s, coinciding with a bond's vibrational period[20].

A very important consideration in a CID experiment is the time needed for the dissociation reaction to take place on the time scale of the experiment. The kinetic window varies widely across different instrument types, from a few tens to hundreds of microseconds for magnetic sector based MS instruments operating at KeV ion kinetic energies, milliseconds in triple-quadrupole and quadrupole-TOF instruments, and from hundreds of milliseconds up to seconds in trapping instruments. For example, the dissociation of large biological ions, with a large number of degrees of freedom, may be relatively slow, and for such systems the kinetic window plays an important role (an excess energy above the thermochemical threshold is needed to observe dissociation). The amount of energy above the critical energy for dissociation required to drive dissociation at a rate that is observable on the time scale of the instrument used is called the kinetic shift and may be very large (up to or even exceeding 100 eV) for large biological ions[21].

The efficiency of the CID process relates to how much of the available precursor ion is converted to product ions, and for MS trapping instruments can be written as[22]

$$[M_p^+]_t = [M_p^+]_0 e^{-nkt} \qquad 3.22$$

Where the $[M_p^+]_t$ is precursor ion abundance after reaction time t, n is the target number density, and k is the rate constant for all precursor ion loss processes. If no precursor ion loss processes other than CID operate, then k represents the rate constant for CID and can be represented as

$$k(E^*) = N_{vr}^{\#}(E^* - E_0)/h\rho_{vr}(E^*) \qquad 3.23$$

where $N_{vr}^{\#}(E^* - E_0)$ is the sum of states at the reaction transition state, ρ_{vr} is the density of states of the excited ion, and E_0 is the threshold energy for the observed dissociation pathway. The dissociation yield depends on the time allowed for the unimolecular reaction to take place and on the number of collisions. The magnitude of variations for the values just described may range widely with the type of instrument used and the design of the experiment, and various values are not independent but influence one another.

Besides the `in trap` hexapole CID in the experiments of this thesis, `nozzle skimmer` CID was often used for generating fragments. Nozzle-skimmer fragmentation refers to the intentional generation of in-source fragmentation by increasing the nozzle-skimmer potential difference (in our setup from tens up to hundreds of volts), typically on instruments with electrospray ion source. When ions are accelerated between the nozzle (cone) and the skimmer (Chapter 2.3.1; Figure 2.17) through a comparatively high pressure region (of a few Torr), CID is efficient. The time scale for this type of dissociation is normally from a few hundred microseconds to a few milliseconds. Both of these CID methods belong to low-energy regime for which some of the main characteristics are the following: the collision energy lies between 10 and 200 eV, the number of collisions is from approximately 10 to 100, the instrument timescale is from a few hundred of microseconds to a millisecond, and approximately 50 % dissociation efficiency can be achieved[22].

3.3. Theoretical aspects

The development of quantum mechanics during the beginning of the 20th century lead to the realization that nearly all properties of chemical systems could in principle be calculated *ab initio*. The development of mathematical and computational methodologies during the middle of the 20th century formed the practical framework for this to be accomplished. Recently, computational resources have reached the level that routine methods are beginning to approach chemical accuracy. The 1998 Nobel Prize in chemistry was awarded to John Pople for the development and use of theoretical chemistry. More recently, in 2013, the Nobel Prize in Chemistry was awarded jointly to Martin Karplus, Michael Levitt and Arieh Warshel for advances in modeling of complex chemical systems. Indeed, the field of computational chemistry is now one of the primary tools of both experimental and theoretical chemists. Today, the highest-level methods can exceed the accuracy attainable in experiment for the smallest systems, while the lower-cost classical or semi-empirical approximations are able to model the intricate details of macro-scale systems containing millions of atoms. However, even with the development of large super-computing centres, in most cases the choice of methodologies remains a trade-off between time, computational resources and required accuracy[23].

Gas phase molecules in a high-vacuum environment can be modelled most accurately by theoretical methodologies on account of their complete isolation and the inherent absence of any possible perturbing influences of surrounding species. For instance, ions trapped in the cell of an FT-ICR MS are under high-vacuum ($\sim 10^{-9}$ torr), and under typical experimental conditions do not experience collisions (mean collision time is on the order of tens of seconds).

3.3.1. Classical molecular mechanics/dynamics

Though not the most accurate, classical molecular mechanics is the most efficient and easily applicable of methods within the zoo of approaches offered by computational chemistry. These methods, based on the laws of classical physics, using force fields to approximate the relative energies of different configurations of a molecule, are applicable to very large systems such as proteins. Many molecular mechanics packages have been devised and applied to a wide range of systems, such as AMBER[24], CHARMM[25] and GROMOS[26].

The molecular mechanics procedure begins with the description of the energy of a system as determined by the force field energy function. These functions contain a set of parameters for each atom-type based on the local environment, including bonded and non-bonded interactions. As such, parameter sets must be defined for atom types in each class of molecular systems, i.e. biomolecules, small organics, inorganic solids, polymers, etc. Typically, parameters are derived from high-level[23] computational methods and/or experimental values. For some specific cases, these approximations are sufficiently accurate, however in most cases, and especially for large systems, they require the consideration of their properties dynamically[27]. Modelling the structure of a protein or peptide, for example, requires the consideration of structural variability associated with secondary structure formation. This necessitates an exploration of the potential energy surface over a large energy/temperature range where many barriers in the system can be crossed[28,29]. First, the system is heated to an elevated temperature so that local barriers are overcome. Then, a procedure of simulated annealing, allows the system to settle back into the various minima on the potential energy surface, after crossing the local barriers, by bringing the temperature/energy of the system back down.

In the work presented in this thesis, initial guess geometries were first optimized in the Gaussian09[30] package using density functional theory (DFT). Charges obtained from these initial results were used for generating parameters of the non-standard terminations of the peptides in their gas-phase conformations using the antechamber program. After an initial molecular mechanics minimization within AMBER, a simulated annealing procedure was used, normally extending up to 1000 K. Within this range, even relatively high barriers, such as *cis/trans* isomerization of peptide bonds, can be overcome. A 1 fs step size was used and coordinates were saved periodically throughout the simulation so that they could then be individually optimized. These locally optimized structures were grouped based on structural similarity and then each further optimized using higher-level methods.

3.3.2. Quantum chemical methods

Classical molecular mechanics models, while relatively simple, are by definition approximations to real systems. Although, the accuracy of the models may be improved, they will fundamentally never truly describe real molecular environments, as they do not consider the quantum nature of matter. A full quantum-mechanical treatment, is exactly the opposite, being by definition the true and complete description of the electronic structure. In the most general sense, this involves solving the Schrödinger equation to obtain the electronic structure of the system. For the smallest systems (i.e. hydrogen) this can be accomplished exactly; however, for virtually all systems, mathematical approximations are required. This section will give a brief description of some of the main concepts and approximations in a quantum-mechanical treatment of molecular systems[29,31-34].

The time-independent Schrödinger equation is an eigenvalue problem that has as its solutions an infinite number of wavefunctions Ψ_n associated with eigenvalues which describe the allowed energy levels, E_n, of the system

$$H\Psi = E\Psi \qquad 3.24$$

Here, the Hamiltonian, H, is an operator describing the total energy of the system of i electrons and a nuclei with masses, m, charge, Z, and separated by distance, r.

$$H = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_a \frac{\hbar^2}{2m_a}\nabla_a^2 - \sum_i\sum_a \frac{e^2 Z_a}{r_{ia}} + \sum_{i<j}\frac{e^2}{r_{ij}} + \sum_{a<b}\frac{e^2 Z_a Z_b}{r_{ab}} \qquad 3.25$$

The first two terms on the right side of the equation are the kinetic energy of the electrons and nuclei; the next term represents the electronic-nuclear attraction and the last two terms describe the electronic and nuclear mutual repulsions. In this description, the wavefunction Ψ depends on the positions of all particles, both electrons and nuclei. The main difficulty in treating real systems arises from the terms containing correlation between particles. This results from the dynamic nature of each particle that is dependent on the properties of every other particle in the system. The first and most important approximation in a typical quantum chemical treatment involves the Born-Oppenheimer approximation, where coupling between the nuclear and electronic motion is neglected. As illustrated in the equation below, invoking this approximation then allows the electronic terms of the Schrödinger equation, H_{elec} and Ψ_{elec}, to be considered for a fixed set of nuclear positions, R_{fixed}, and the repulsive nuclear-nuclear potential calculated separately as a constant. This assumption is normally very reliable due to the scale of the mass difference ($m_{electron} = 9.1 \times 10^{-31}$ kg and $m_{proton} = 1.7 \times 10^{-27}$ kg).

$$(H_{elec} + V_{Nuc})\Psi_{elec}(r; R_{fixed}) = E_{elec}\Psi_{elec}(r; R_{fixed}) \qquad 3.26$$

3.3.3. Hartree Fock theory and density functional theory

Hartree-Fock theory (HF) extends the approximation to the electron-electron correlation in the description of the electronic structure of the system. By assuming that each electron sees all other electrons as a static field, it is possible to solve a series of one-electron problems. This strategy avoids the most difficult part of the electronic Schrödinger equation, namely electron correlation. While approximately valid for some cases, neglecting electron-electron correlation comes with some significant risks, especially when dealing with systems heavily dependent on the description of dispersion, such as systems having a significant extent of hydrogen bonding[35,36]. Many post-HF methodologies include, to different extents, electron-electron correlation. However, these methods are generally very costly. Possibly the most common and least computationally demanding of these methods is Møller–Plesset (MP) perturbation theory. Second order MP theory (MP2) can be used to calculate geometries and frequencies of small systems, and electronic energies of moderately sized systems and accounts for a large amount of the correlation energy. Higher-order MP theory and other post-HF methods, while very accurate (such as coupled cluster (CC) theory) are exceedingly costly for most practical uses.

An alternative to these so called wavefunction-based methods is Density functional theory (DFT). Practically speaking, DFT is appealing in that much of the electron-electron correlation energy is implicitly included in the calculation, without requiring the large computational cost of most post-HF methods. In fact, for a variety of cases, DFT is now the method of choice and in many situations, when properly applied, can be of similar or even better accuracy than MP2 theory (for example, in the calculation of vibrational frequencies). Due to the appealing balance between accuracy and computation cost, DFT, and specifically the B3LYP functional, is the only quantum chemical approach used in the work described in this thesis.

DFT was establish by the publication of the Hohenberg and Kohn theorems[37]. They demonstrated that the ground-state electron density $\rho(r)$ uniquely relates to the external potential (the nuclei) and, as a consequence, the Hamiltonian and wavefunction. Kohn and Sham later developed a procedure using a Hamiltonian for a fictitious system of non-interacting electrons[38]. In this specific case, the Hamiltonian becomes separable and can be treated as a sum of one-electron operators. The functional of the energy is described by the following expression.

$$E[\rho(r)] = T_{ni}[\rho(r)] + V_{ne}[\rho(r)] + V_{ee}[\rho(r)] + \Delta T[\rho(r)] + \Delta V_{ee}[\rho(r)] \qquad 3.27$$

Here, $T_{ni}[\rho(r)]$ is the kinetic energy term for the non-interacting electrons and $V_{ne}[\rho(r)]$ and $V_{ee}[\rho(r)]$ contain the Coulomb potential energy of the nuclear-electronic and electronic-electronic interactions. Following this are the corrections to this treatment, mainly the kinetic energy difference between the corresponding system of interacting electrons and all the corrections to the electron-electron interactions arising from the quantum nature of the system. These corrections are the commonly referred to exchange-correlation (XC) energy.

One of the most commonly used functionals today is the B3LYP functional, containing the LYP correlation functional and the B3 (3-parameter) form of Becke's exchange functional[39-42]. This is a hybrid functional with the exchange of the non-interacting system calculated within the Hartree-Fock formalism. This accounts for a major component of the exchange energy. The corrective exchange-correlation portion of the B3LYP functional is summarized below (a = 0.20, b = 0.72, c = 0.81). Here, the superscript $LSDA$ refers to the local spin density approximation, HF to the Hartree-Fock calculated portion of the exchange energy and x and c to the exchange and correlation terms, respectively.

$$E_{xc}^{B3LYP} = (1-a)E_x^{LSDA} + aE_x^{HF} + bE_x^{B} + (1-c)E_c^{LSDA} + cE_c^{LYP} \qquad 3.28$$

3.3.4. Basis sets

The wavefunction of the system is modelled using a set of mathematical basis functions. The better these functions are in terms of accuracy and flexibility, the more complete the description of the electronic structure will be. However, the cost of the calculation scales sharply as a function of the size of the system (number of electrons) and the number (N) of basis functions (for most DFT methods ~N^3, but for some post-HF methods >N^7). For this reason, it is necessary to use basis functions that are both accurate physically, but can also be efficiently computed and are mathematically flexible. One of the initial advances in modern computational chemistry was the treatment of the electronic wavefunction, Φ, as a linear combinations of atomic orbitals (LCAO), φ_i.

$$\phi = \sum_{i=1}^{N} a_i \varphi_i \qquad 3.29$$

A set of N functions, φ_i, thus serve as the basis set and together with the coefficients a_i are used to construct a mathematically flexible wavefunction. The use of a larger basis set simply makes the mathematical description of the wavefunction more complete, but comes at an increased computational cost. For reasons of computational efficiency, hydrogen-like wavefunctions (Slater type orbitals) are often replaced by Gaussian functions to model the atomic orbitals in LCAO wavefunctions[43].

$$\phi(x, y, z; \alpha, i, j, k) \sim x^i y^j z^k e^{-\alpha(x^2+y^2+z^2)} \qquad 3.30$$

Here, parameters *(i, j, k)* determine the orbital type. The spherical *s*-type orbital has parameters *(i, j, k) = 0* and the p-orbital has one *i, j* or *k* =1. Three possible permutations in constructing p-orbitals exist from *i=1, j=1* or *k=1* giving the p_x, p_y, p_z orbitals. When i+ *j*+ *k* = 2, a d-type function results, and this scheme can be extrapolated to higher angular momentum atomic orbitals.

Various problems associated with the mathematical character of Gaussian functions (unphysical behavior) are accounted for by using linear combinations of Gaussians, rather than single functions.

$$\varphi(x, y, z; \{\alpha\}, i, j, k) = \sum_{\alpha=1}^{M} c_\alpha \phi(x, y, z; \alpha_\alpha, i, j, k) \qquad 3.31$$

The coefficient c_α optimizes the contribution of each individual function and allows for normalization. In all currently employed Gaussian-based basis sets, multiple sums of Gaussian functions are used to model the orbitals in a molecule. Core orbitals are often described with a single set of Gaussian functions, being less dependent on the local bonding environment. In the 6-31G basis set, the core orbitals are represented with six Gaussian functions and the valence orbitals are modelled with both a set of three functions and a single function. Polarization functions add spatial flexibility by including atomic functions of higher angular momentum than the occupied orbitals. As an illustration, the 6-31G(d,p) basis set adds d-type functions on non-hydrogen atoms and p-type functions on hydrogen atoms. For similar purposes, so called diffuse functions, spatially larger versions of the basis functions, are added to all non-hydrogen atoms (+) or all atoms (++) when additional flexibility to the orbital size is thought to be necessary, such as for anions.

All of the calculations reported in this thesis used the 6-31++G(d,p) basis set, as it was found to be both computationally efficient and sufficiently accurate when considering peptides of small to moderate size (including structures, electronic energies, and vibrational and thermochemical properties).

3.3.5. Vibrational analysis and thermodynamics

After an optimized geometry for a molecular structure is obtained, other properties of the system can be evaluated. Of particular interest in the work described in this thesis is the calculation of vibrational frequencies and free energies. For vibrational analysis, it is necessary to consider the Hessian matrix, a 3*N* x 3*N* matrix containing the second partial derivatives of the potential energy as a function of the nuclear coordinates.[32,33] This provides the vibrational normal coordinates and the mass-weighted force constants related to the vibrational frequencies. As well, calculated intensities are obtained by relating the derivative of the dipole moment to the normal coordinates. If a geometry corresponding to a minimum on the potential energy surface is located, all vibrational frequencies should be non-imaginary and non-zero. In the case that a single imaginary frequency is found, this indicates a transition state structure, while more than one imaginary frequency indicates a higher-order saddle-point structure. Generally, errors in vibrational frequencies associated with usage of the harmonic oscillator approximation and with basis set effects, among other approximations, are adequately corrected for using small, linear scaling factors[44,45] (for example, 0.975 is used here for the B3LYP/6-31++G(d,p) level of theory).

Many electronic structure packages (including Gaussian09) incorporate programs to calculate thermodynamic properties (free energies, entropies, etc.) of a molecular system. Using Boltzmann statistics and the calculated Gibbs free energies, one can also predict the relative abundances, at temperature T, within a set of conformational isomers N_B and N_A.

$$K = \frac{N_B}{N_A} = e^{\frac{-\Delta G_T^{\circ}}{RT}} \qquad 3.37$$

Assuming a thermal distribution of internal energies, isomers connected by a barrier exist in abundances corresponding to their free energies. For instance, if isomer *A* is 5 kJ mol^{-1} higher in free energy than isomer *B,* they occur in a 7.5:1 ratio while for a 20 kJ mol^{-1} difference, the ratio is 3200:1. All discussions of thermodynamic values that follow refer to values calculated at 298.1 K.

The practical application of these general theoretical approaches to the studies reported in this thesis are described in more detail in each of the individual chapters.

REFERENCES

(1) Herschel, W. *Experiments on the refrangibility of the invisible rays of the Sun* Phil. Trans. R. Soc. Lond. **1800**, *90*, 284.

(2) Coblentz, W. W. *Investigations of Infra-Red Spectra*; Carnegie institution of Washington, **1905.**

(3) Stearns, J. A.; Seaiby, C.; Boyarkin, O. V.; Rizzo, T. R. *Spectroscopy and conformational preferences of gas-phase helices* Phys. Chem. Chem. Phys. **2009**, *11*, 125.

(4) Colthup, N. B.; Daly, L. H.; Wiberley, S. E. *Introduction to infrared and Raman spectroscopy*; Academic Press, **1990**.

(5) Shin, S. K.; Beauchamp, J. L. *Infrared multiphoton dissociation spectrum of* $CF_3Mn(CO)_3(NO)^-$ J. Am. Chem. Soc. **1990**, *112*.

(6) Dunbar, R. C. *Photodissociation of the methyl chloride* (CH_3Cl^+) *and nitrous oxide* (N_2O^+) *cations* J. Am. Chem. Soc. **1971**, *93*, 4354.

(7) Freiser, B. S.; Beauchamp, J. L. *Gas phase ion chemistry and photochemistry of ions generated from perfluoropropylene. Photodissociation of the perfluoroallyl cation* J. Am. Chem. Soc. **1974**, *96*, 6260.

(8) Okumura, M.; Yeh, L. I.; Myers, J. D.; Lee, Y. T. *Infrared spectra of the solvated hydronium ion: Vibrational predissociation spectroscopy of mass-selected* $H_3O^{+}\cdot(H2O)_n\cdot(H_2)_m$ J. Phys. Chem. **1990**, *94*, 3416.

(9) Watson, C. H.; Zimmerman, J. A.; Bruce, J. E.; Eyler, J. R. *Resonance-enhanced two-laser infrared multiple photon dissociation of gaseous ions* J. Phys. Chem. **1991**, *95*, 6081.

(10) Peiris, D. M.; Cheeseman, M. A.; Ramanathan, R.; Eyler, J. R. *Infrared multiple photon dissociation spectra of gaseous ions* J. Phys. Chem. **1993**, *97*, 7839.

(11) V.N. Bagratashvili; V.S. Letokhov; A.A. Makarov; Ryabov, E. A. *Multiple Photon Infrared Laser Photophysics and Photochemistry*; Gordon and Breach: New York, **1985.**

(12) Makarov, G. N.; Petin, A. N. *Isotopically selective IR multiphoton dissociation of* SF_6 *in a pulsed gas dynamic flow interacting with solid surface* Chem. Phys. **2001**, *266*, 125.

(13) Black, J. G.; Yablonovitch, E.; Bloembergen, N.; Mukamel, S. *Collisionless Multiphoton Dissociation of* SF_6*: A Statistical Thermodynamic Process* Phys. Rev. Lett. **1977**, *38*, 1131.

(14) Thompson, J. J. *Rays of positive electricity* Proceedings of the Royal Society **1913**, *A 89*, 1.

(15) Jennings, K. R. *Collision-induced decompositions of aromatic molecular ions* Int. J. Mass Spectrom. Ion Proc. **1968**, *1*, 227.

(16) McLafferty, F. W.; Fairweather, R. B. *Metastable ion characteristics. VIII. Characterization of ion decomposition mechanism by metastable ion abundances* J. Am. Chem. Soc. **1968**, *90*, 5915.

(17) Drahos, L.; Vekey, K. *MassKinetics: a theoretical model of mass spectra incorporating physical processes, reaction kinetics and mathematical descriptions* J. Mass Spectrom. **2001**, *36*, 237.

(18) Douglas, D. J.; French, J. B. *Collisional focusing effects in radio frequency quadrupoles* J. Am. Soc. Mass Spectrom. **1992**, *3*, 398.

(19) Goeringer, D. E.; McLuckey, S. A. *Evolution of ion internal energy during collisional excitation in the Paul ion trap: A stochastic approach* J. Chem. Phys. **1996**, *104*, 2214.

(20) Schwartz, R. N.; Slawsky, Z. I.; Herzfeld, K. F. *Calculation of Vibrational Relaxation Times in Gases* J. Chem. Phys. **1952**, *20*, 1591.

(21) Lifshitz, C. *Kinetic shifts* Eur. J. Mass Spectrom. **2002**, *8*, 85.

(22) Mitchell Wells, J.; McLuckey, S. A. In *Methods Enzymol*; Burlingame, A. L., Ed.; Academic Press: 2005; Vol. Volume 402, p 148.

(23) Friesner, R. A. *Ab initio quantum chemistry: Methodology and applications* Proc. Natl. Acad. Sci. **2005**, *102*, 6648.

(24) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *The Amber biomolecular simulation programs* J. Comput. Chem. **2005**, *26*, 1668.

(25) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *CHARMM: A program for macromolecular energy, minimization, and dynamics calculations* J. Comput. Chem. **1983**, *4*, 187.

(26) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *GROMACS: Fast, flexible, and free* J. Comput. Chem. **2005**, *26*, 1701.

(27) Karplus, M.; Kuriyan, J. *Molecular dynamics and protein function* Proc. Natl. Acad. Sci. **2005**, *102*, 6679.

(28) Adcock, S. A.; McCammon, J. A. *Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins* Chem. Rev. **2006**, *106*, 1589.

(29) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models* **2004**.

(30) M. J. Frisch, G. W. T., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, *Gaussian 09, Revision A.1* **2009**.

(31) Jensen, F. *Introduction to Computational Chemistry* **2007**.

(32) Szabo, A.; Ostlund, N. *Modern Quantum Chemistry - Introduction to Advanced Electronic Structure Theory* **1996**.

(33) Levine, I. N. *Quantum Chemistry* **2005**.

(34) Roothaan, C. C. J. *New Developments in Molecular Orbital Theory* Rev. Mod. Phys. **1951**, *23*, 69.

(35) Schwabe, T.; Grimme, S. *Double-hybrid density functionals with long-range dispersion corrections: higher accuracy and extended applicability* Phys. Chem. Chem. Phys. **2007**, *9*, 3397.

(36) Grimme, S. *Accurate description of van der Waals complexes by density functional theory including empirical corrections* J. Comput. Chem. **2004**, *25*, 1463.

(37) Hohenberg, P.; Kohn, W. *Inhomogeneous Electron Gas* Phys. Rev. **1964**, *136*, B864.

(38) Kohn, W.; Sham, L. J. *Self-Consistent Equations Including Exchange and Correlation Effects* Phys. Rev. **1965**, *140*, A1133.

(39) Becke, A. D. *Density functional thermochemistry. III. The role of exact exchange* J. Chem. Phys. **1993**, *98*, 5648.

(40) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Results obtained with the correlation energy density functionals of becke and Lee, Yang and Parr* Chem. Phys. Lett. **1989**, *157*, 200

(41) Becke, A. D. *Density-functional exchange-energy approximation with correct asymptotic behavior* Phys. Rev. A **1988**, *38*, 3098.

(42) Lee, C.; Yang, W.; Parr, R. G. *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density* Phys. Rev. B **1988**, *37*, 785.

(43) Hehre, W. J.; Stewart, R. F.; Pople, J. A. *Self Consistent Molecular Orbital Methods. I. Use of Gaussian Expansions of Slater Type Atomic Orbitals* J. Chem. Phys. **1969**, *51*, 2657.

(44) Alecu, I. M.; Zheng, J.; Zhao, Y.; Truhlar, D. G. *Computational Thermochemistry: Scale Factor Databases and Scale Factors for Vibrational Frequencies Obtained from Electronic Model Chemistries* J. Chem. Theory Comput. **2010**, *6*, 2872.

(45) Andersson, M. P.; Uvdal, P. *New Scale Factors for Harmonic Vibrational Frequencies Using the B3LYP Density Functional Method with the Triple-ζ Basis Set 6-311+G(d,p)* J. Phys. Chem. A **2005**, *109*, 2937.

Chapter 4.

Structure of anionic c-type peptide fragments elucidated by IRMPD spectroscopy

In contrast to protonated peptides, collision-induced dissociation of deprotonated peptides is known to yield abundant *c*- and *z*-type fragments in addition to typical *a*-, *b*- and *y*-type ions. Here we investigate the isomeric structures of short anionic *c*-type peptide fragments using IR photodissociation spectroscopy. As for all N-terminal fragments of deprotonated peptides, an important structural question concerns the site of deprotonation, as no C-terminal COOH group is present in these products. Comparison of the experimental action spectra with spectra computed for several candidate structures suggests that the c_1 and c_2 ions investigated have linear peptide structures with a C-terminal -C(=O)-NH_2 amide group. Competition between deprotonation on the peptide bond nitrogen atom, forming an amidate anion, and on the residue side chain, depending on its gas-phase acidity, is observed. *

---

*Adapted from: J.Grzetic and J.Oomens, *Structure of anionic c-type peptide fragments elucidated by IRMPD spectroscopy*, Int. J. of Mass Spectrom.,(2012), 316-318, 216-226

4.1. Introduction

Collision induced dissociation (CID) tandem mass spectrometry forms one of the cornerstones of modern proteomics research as it enables the routine determination of the sequence of amino acid residues in a peptide[1-4]. Despite its widespread practical application, a molecular-level understanding of the actual fragmentation reactions occurring in CID has only been partly established and remains a subject of much debate. Prof. Alex Harrison has been one of the driving forces in understanding the fragmentation chemistry of both protonated and deprotonated peptides[5-10]. Quite obviously, accurate knowledge of the molecular structures of the reactants and products in peptide fragmentation reactions greatly aids in modeling the reaction pathways[11]. While fragment structures have long been inferred from the m/z values in MS^n experiments, the combination of tunable infrared laser sources with MS/MS instrumentation has rapidly revolutionized peptide fragment structure identification[12-14]. Infrared multiple photon dissociation (IRMPD) spectra of mass selected ions in combination with quantum-chemically computed spectra for candidate structures have proven to be very decisive in the identification of specific fragment structures[15,16]. Results of these experiments, possibly in combination with calculations of intermediates and transition states[17] as well as with experimental determination of energetic and entropic barriers (*e.g.* by threshold CID[18]) are expected to eventually lead to a more complete picture of peptide fragmentation pathways.

Spectroscopic studies of CID fragments derived from protonated peptides have focused mainly on *a*- and *b*-type ions. It was for instance shown that the b_2 ion from Ala_3 and AlaGlyGly forms an oxazolone structure[19,20], while b_2 from $HisAla_4$ is a diketopiperazine ion[21]. For larger *b*-type fragments 'head-to-tail' cyclization producing macrocyclic structures has been spectroscopically identified, which raises the important question of whether it may induce sequence scrambling in MS^n studies [6,7,22-24]. Similar formation of macrocyclic structures possibly leading to a rearrangement of the linear sequence was recently reported for *a*-type peptide fragment ions as well[25,26].

Although most MS instruments can operate in both positive and negative ion mode, peptide sequencing is customarily performed on the protonated species; sequencing in negative ion mode is possible, however, and some applications have been reported, see e.g. Ref[27]. Quite obviously, sequence coverage could be increased by analyzing the same peptide in both positive and negative ion mode and using complementary information[28]. Nonetheless, information on the deprotonated species is hardly used in practical proteomics applications, which is perhaps partly due to the fragmentation processes being not as extensively studied. Whereas backbone fragmentation of protonated peptides results mainly in *a-, b-* and *y*-type ions, negative ion CID mass spectra have been reported to feature in addition *c* and *z-type* fragment ions. *c*-type ions have been observed in CID tandem mass spectra of anionic peptide species containing alkyl[29], acidic[10,30,31], hydroxyl [32], basic[33] and aromatic[5,34] residues, sometimes even with very high abundance. These *c-type* fragments generated from anionic peptides provide information that is not available from their positive-ion counterparts, and their identification can be used to determine the sequence from the N-terminus and in that way increase sequence coverage.

Many studies have addressed the fragmentation of deprotonated peptides based on MS/MS data[5,29-32,35-40], while more recently IR spectroscopy has also been applied to elucidate the structure of anionic peptide backbone fragments. IR spectroscopy showed the a_3^- fragment of deprotonated Ala_3 possesses a linear amidate structure, where the charge is formally localized on one of the amide nitrogen atoms[41]. Because dissociation reactions are generally charge directed, determination of the deprotonation site is of importance in understanding the fragmentation pathways. Very recently, the b_2 fragments generated from deprotonated Ala_3 and deprotonated AlaTyrAla have been spectroscopically shown to possess oxazolone structures [42], as had indeed been suggested by Harrison[9,36]. Here we use IRMPD spectroscopy to establish the structures of short anionic *c*-type fragments.

Finally, it is interesting to note that while *c*-type ions generally do not occur in CID tandem mass spectra of protonated peptides, there has been substantial interest in the structure of positively charged *c-type* fragment ions, as they constitute the main dissociation products in electron capture and electron transfer dissociation (ECD/ETD) methods, which are increasingly applied in top-down proteomics. In ECD/ETD, multiply protonated peptides (or entire proteins) are partially reduced through the capture of an electron thus producing an intermediate radical cation, which then fragments mainly by backbone N-C_α bond cleavage yielding *c*- and *z*-type ions[43-47]. Mechanistic studies of these processes[48-50] have often speculated on the structure of the resulting *c*-type ion as having either an amide or an enol-imide structure, which may give information on the reaction pathways of ECD/ETD. An IRMPD spectroscopic study on a model *c*-type ion generated by ECD established an amide structure, although post-dissociation isomerization could not be excluded[51].

Here we report the first IRMPD spectra for anionic *c*-type ions, in particular the c_1-ions generated from deprotonated AlaPhe, TyrPhe, and TrpPhe, as well as the c_2-ions generated from deprotonated Leucine-Ala^2-Enkephaline (YAGFL) and GlyGlyPhe.

4.2. Experimental and computational methods

4.2.1. IRMPD spectroscopy

Infrared spectra are obtained by infrared multiple-photon dissociation (IRMPD) spectroscopy using a home-built Fourier Transform ion cyclotron resonance mass spectrometer (FTICR-MS) equipped with a 4.7 T actively-shielded superconducting magnet

(Cryomagnetics Inc., Oak Ridge, TN, USA) coupled to the beamline of the infrared free-electron laser FELIX as described in detail elsewhere[13,52].

Deprotonated peptide anions are generated by electrospray ionization (Micromass Z-Spray source), which facilitates gentle ionization and vaporization of labile biomolecular species. Working solutions contain about 1 mM of the peptide in a 7:3 mixture of methanol and water with 1 mM NaOH added to enhance deprotonation of the peptides. Sample peptides Leu-Ala^2-Enkephaline (YAGFL), H-Tyr-NH_2, Gly-Gly-NH_2, AlaPhe and GlyGlyPhe are obtained from Bachem (Weil am Rhein, Germany) and TyrPhe is obtained from Sigma-Aldrich (Zwijndrecht, The Netherlands). CID fragments from GlyGlyPhe are generated by nozzle-skimmer dissociation in the high-pressure region at the inlet of the mass spectrometer. Ions are accumulated in a linear hexapole trap before they are transferred to the ICR cell, where the anion of interest (c_1 or c_2) is isolated using a SWIFT excitation pulse[53]. Dissociation of TrpPhe, TyrPhe, AlaPhe and YAGFL is induced in the ICR-cell by a 0.5 s long CO_2 laser pulse (Universal laser systems, Scottsdale, USA, ULR-25). The choice between the two dissociation methods (front end CID and CO_2-laser IRMPD) depended on which of the two was more efficient; for the systems where both fragmentation methods were used, no differences were found in the fragment IR spectra.

Mass-isolated ions are irradiated with the tunable IR light from the free electron laser FELIX[54]. FELIX produces 5 μs long pulses with an energy of ~60 mJ (~30-40 mJ in the ICR cell) and the full width half maximum bandwidth of ~0.5% of the central wavelength. In the current experiments, the FELIX wavelength is scanned from 5.5 to 13 μm, which corresponds to a spectral range of 800-1800 cm^{-1}. Ions are irradiated for 3.8 s, while the entire FTICR sequence is 6 s long.

Infrared multiple-photon dissociation (IR-MPD) spectroscopy is based on the principle that only photons of energy resonant with a vibrational transition are absorbed. Upon resonant photon absorption, the vibrational energy of the ion is rapidly quenched by other internal degrees of freedom through intramolecular vibrational redistribution (IVR); tens to hundreds of photons are typically absorbed leading to dissociation of the anion[12,55-57]. After irradiation, a mass spectrum is generated using an excite/detect sequence as implemented in the FTICR control software, which is a modified version of that described by Heeren and co-workers[58]. Three to five mass spectra are averaged at each wavelength point. Recording the photofragment yield as a function of wavelength gives an IR spectrum, where the yield is determined as ∑I(*fragments*) / ∑I(*all ions*). The yield is corrected for the laser power profile over the scan range assuming a linear dependence of the fragmentation yield on laser power.

4.2.2. Density functional theory calculations

Calculations are performed using the B3LYP functional and the 6-31++G(d,p) basis set as implemented in Gaussian03. This method was previously found to accurately reproduce relative energies and IR spectra for deprotonated species[59,60]. Starting structures are generated using chemical intuition, where the number of hydrogen bonding interactions is maximized and eclipsed bond interactions are avoided. Also, only *trans*-configurations of the peptide bonds are considered. Experimental spectra are compared to calculated spectra to identify the molecular structures present in the mass spectrometer. Computed harmonic frequencies are scaled by 0.97 as has been found appropriate for this method and basis set in this wavelength range and convoluted with a 25 cm^{-1} fwhm Gaussian line shape function to facilitate easy comparison with experimental spectra.

4.3. Results

c-type ions are N-terminal fragments resulting from cleavage of the N-C_α bond, which have generally been suggested to be linear peptide fragments with an $-C(=O)-NH_2$ amide moiety at the C-terminus. Deprotonation of a peptide no longer possessing a C-terminal carboxylic acid group raises the immediate question of where the charge (formally) resides. Figure 4.1 lists suggested linear c_2 structures with different possible deprotonation sites. Structures considered as likely candidates are especially those stabilized by conjugation, thus delocalizing the negative charge providing increased gas-phase stability, and with maximized number of hydrogen bonding interactions (which are obviously not included in Figure 4.1). Deprotonation of the C-terminal amide group (structure I) results in an amidate resonance, similar to deprotonation of the peptide bond nitrogen. Anionic *a*-type fragments without acidic residues were shown to possess such a deprotonated peptide bond, forming an amidate structure[41,61]. Structure III, having a deprotonated N-terminal amino group, did not converge in the calculations, probably because no reasonable delocalization of the negative charge is possible. Structures IV and V are enolate resonances with the C-terminal amide carbonyl and the peptide bond carbonyl, respectively. Structure VI represents fragment ions with a deprotonated side chain, which is obviously only possible for selected amino acid residues, e.g. these structures are unlikely for alkyl side chains. Finally, structure VII shows the tautomeric enol-imide form of the C-terminal amide.

For c_1-type ions, analogous structures are suggested, with the obvious exception that deprotonation of the peptide bond is not possible, hence excluding the amidate and enolate structures II and V, respectively. Using the results from ion spectroscopy, we will here investigate the structure for the simple c_1 and c_2 ions generated from AlaPhe and GlyGlyPhe, respectively, and then investigate the influence of selected functionalized side chains.

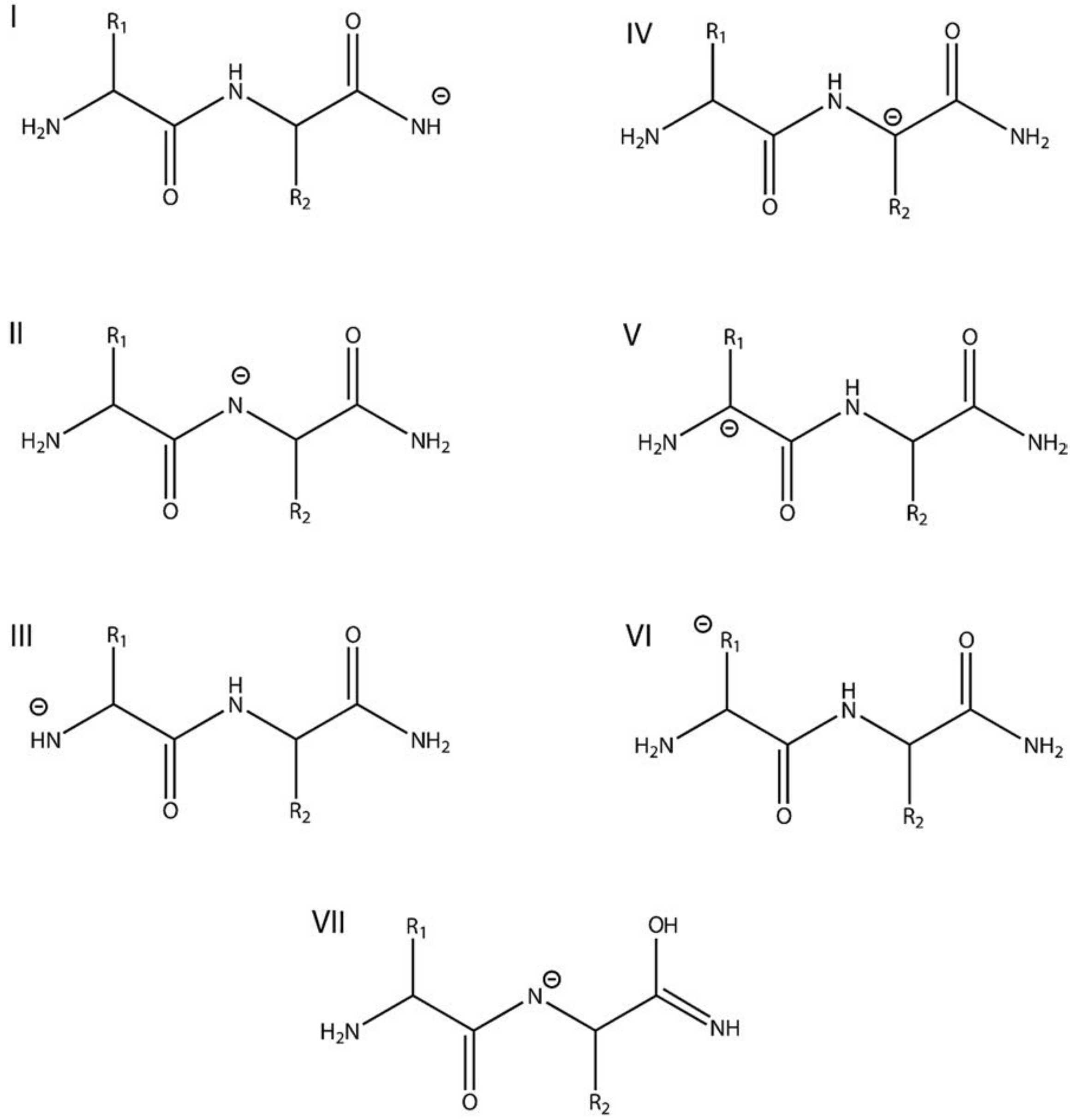


Figure 4.1. Suggested structures and deprotonation sites for *c-type* fragments. Structure VI represents a side-chain deprotonated c-ion, of which the precise structure obviously depends on the residue.

4.3.1. c_1-anion from AlaPhe

The c_1 anion from deprotonated AlaPhe (m/z 87) was mass-isolated and upon IRMPD found to undergo dissociation into four mass channels, being in order of their abundances m/z 70 (loss of ammonia), m/z 26 (cyanide anion), m/z 42 (NCO^-) and m/z 44 (NH=CH-O^-).

In Figure 4.2 the experimental spectrum is compared with calculated spectra for selected candidate structures. Calculated relative free energies are also given for each of the structures. All structures result from a linear peptide backbone structure with a C-terminal amide moiety, where structure **a** is deprotonated at the C-terminal amide NH_2 (analogous to structure I in Figure 4.1. Structure **b** is an enolate anion resulting from deprotonation of the α-carbon atom (IV), while **c** is deprotonated on the N-terminal amino group (III).

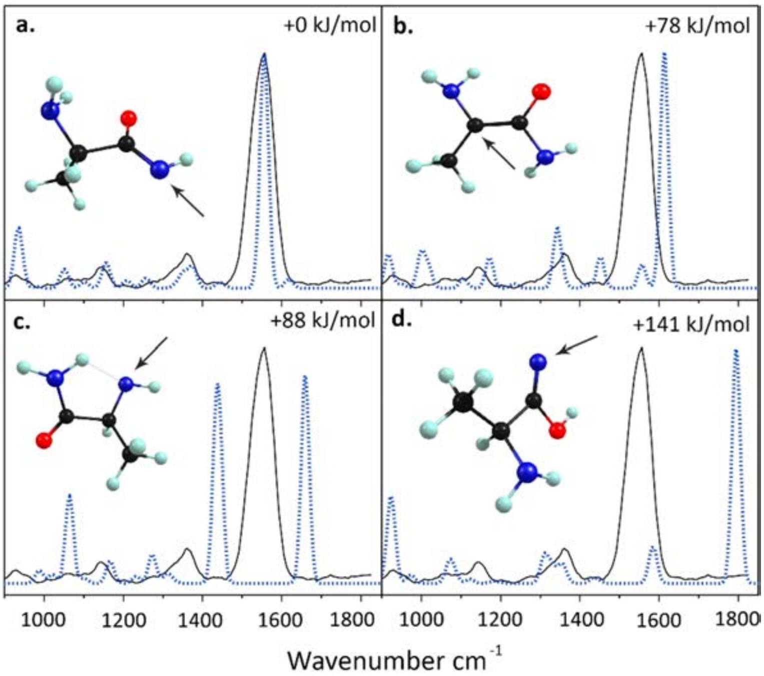


Figure 4.2. Experimental IRMPD spectrum (solid black) of the *c_1* anion from deprotonated AlaPhe compared with calculated spectra (dotted lines) for structures deprotonated at C-terminal amide moiety (**a**), the α-C atom (**b**), the N-terminal amino group (**c**) and at the C-terminal amide with enol-imine tautomerization (**d**). Calculated relative free energy values (in kJ mol^{-1}) are given for each of the isomers. The arrow indicates the formal deprotonation site, although the charge is delocalized due to conjugation.

Structure **d** is the enol-imide tautomer of structure **a**. Only the calculated spectrum for structure **a** is in agreement with the experimental one, while spectra for the other candidate structures clearly do not match. Also there is a significant calculated energy difference between structure **a** and the other structures. The most intense band in the spectrum is due to the carbonyl stretching mode at 1550 cm^{-1}. Another feature of special interest is the band between 1300 and 1400 cm^{-1}, which corresponds to the N-H bending mode of the deprotonated amide moiety. This feature compares favorably with the experimental spectrum of deprotonated acetamide, which was shown to adopt an amidate rather than an enolate form in the gas phase[41], providing further support for the assigned c_1-ion structure.

4.3.2. c_2-anion from GlyGlyPhe

The c_2 anion from deprotonated GlyGlyPhe (m/z 130) was mass-isolated and upon IRMPD found to undergo dissociation into only one observable mass channel, m/z 112, corresponding to water loss. Loss of water from c_2-type fragment anions was also observed by Harrison in CID tandem MS experiments[5]. The experimental IRMPD spectrum compared with theoretical spectra for selected structures from 1 is presented in Figure 4.3. Calculated relative free energies at 298 K for each of the structures are also given in Figure 4.3. Panels **a** through **e** show calculated spectra and structures for the peptide bond amidate isomer (II in Figure 4.1), the deprotonated C-terminal amide isomer (I), the amidate isomer with tautomeric enol-imide motif (VII) and enolate isomers resulting from deprotonation on the α-carbon atoms of the C-terminal Gly residue (IV) and of the N-terminal Gly-residue (V), respectively.

Panel **a** shows in addition a scan at reduced laser power, which accurately reproduces the spectrum of the c_2 anion confirming the double-peak structure of the features centered around 1350 and 1525 cm^{-1}, so that there are five main bands between 1200 and 1800 cm^{-1}.

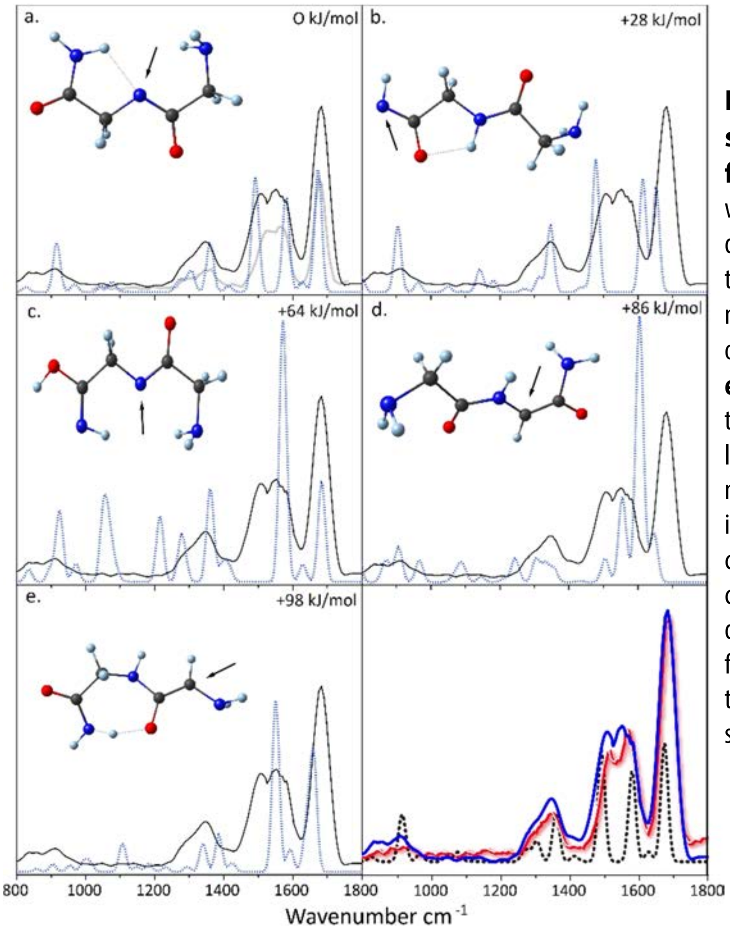


Figure 4.3. Experimental IRMPD spectrum (solid black) of the *c_2* anion from deprotonated GlyGlyPhe compared with calculated spectra (dotted lines) for structures deprotonated on the central amide nitrogen atom (**a**), the C-terminal amide nitrogen (**b**), the central amide nitrogen but with enol-imide C-terminal group (**c**) and on the C-terminal and N-terminal α-carbon atom (**d** and **e**, respectively). The grey trace in panel **a** corresponds to the experimental spectrum recorded at 50% reduced laser power. Calculated relative free energy values (in kJ mol^{-1}) are given for each of the isomers. The arrow indicates the formal deprotonation site, although the charge is delocalized due to conjugation. Panel **f** compares the experimental IRMPD spectra of deprotonated GlyGlyNH$_2$ (blue, thick) with that of the c_2 fragment from deprotonated GlyGlyPhe (red, thin). The theoretical spectrum (black, dotted) for the amidate structure II as displayed in Figure 4.1 is also shown.

Inspection of the five calculated spectra in Figure 4.3 and comparing them to the experimental spectrum clearly suggests that only the amidate structure in panel **a**, resulting from deprotonation of the first peptide amide bond, can reasonably account for these five absorption bands. This structure corresponds to the thermodynamically most favorable isomer of those displayed in Figure 4.3. The alternative amidate structure in panel **b**, resulting from deprotonation of the C-terminal amide moiety and higher in free energy by 28 kJ mol^{-1}, presents a very different calculated spectrum, clearly deviating from the experimental band positions. The main feature in the experimental spectrum at 1680 cm^{-1} is due to the C=O stretch mode of the non-conjugated carbonyl group. The band at 1554 cm^{-1} is due to CO stretching of the conjugated carbonyl group (*i.e.* adjacent to the deprotonation site) and is therefore diagnostic of species having an amidate or enolate structure. This feature is analogous to that observed for the amidate a_3^- fragment ion from deprotonated Ala_3[41], where the calculation also predicts a value for the asymmetric stretching mode of the conjugated N-C-O group that is slightly too high. Furthermore, the absorption at 1360 cm^{-1} is due to C-C stretching with concomitant scissoring of the N-C-O moiety, which is also very characteristic for the amidate structure. Finally, the relatively strong feature at 1505 cm^{-1} is due to an NH_2 scissoring mode

To further support our assignment of the c_2 anion structure as a linear peptide fragment with an amide moiety at the C-terminus, an IRMPD spectrum was recorded for the conjugate base of the commercially available diglycinamide (GlyGly-NH_2). Panel **f** of Figure 4.3 shows the IRMPD spectrum of deprotonated GlyGly-NH_2 overlaid on the experimental spectrum of the c_2 anion of GlyGlyPhe. Within the experimental accuracy, the two spectra are identical and moreover, the same dissociation channels are observed upon IRMPD. Hence, we conclude that the c_2 anion has a linear peptide structure with a C-terminal amide moiety, deprotonated on the amide bond linking the two Gly residues, forming amidate structure II (Figure 4.1). This structure corresponds to the lowest-energy isomer and formation of an amidate motif is analogous to that observed for *a*-type fragment anions without functionalized side chains[41,61].

4.3.3. TyrPhe c_1-anion

The c_2-anion containing only Gly residues forms a linear C-terminal amide structure deprotonated on the amide nitrogen of the peptide bond connecting the two Gly residues. Other deprotonation motifs of the backbone, such as deprotonation of the C-terminal amide or of one of the α-carbon atoms, leading to enolate structures, are substantially higher in energy and can be excluded based on features inconsistent with the experimental IR spectrum. Deprotonation of the C-terminal amide group is only favored when no peptide bond amide linkage is available, such as in the case of the c_1 anion (Figure 4.2). We now address the question whether side-chain deprotonation can compete with backbone deprotonation investigating the influence of two moderately acidic residues, Tyr and Trp, on the *c*-fragment ion structure.

Upon CO$_2$ laser irradiation, deprotonated TyrPhe at m/z 327 was observed to fragment mainly into c_1 (m/z 179), z_1 (m/z 147) and y_1 (m/z 164) fragments. Formation of both c_1 and z_1 ions may indicate that the dissociation mechanism involves proton transfer from the peptide bond nitrogen to the deprotonated carboxylate moiety, forming an amidate intermediate, as was suggested by Harrison [5]. Upon cleavage of the N-C$_\alpha$ bond, the charge can remain on the N- or C-terminal fragment resulting in *c*- or *z*-ions. Upon IRMPD, the isolated c_1- fragment was found to undergo dissociation forming fragments with m/z 107, 71 and 106, in order of their abundances. Ions at m/z 107 and 106 probably correspond to the Tyr side chain CH$_3$-C$_6$H$_4$-O$^-$ and its corresponding radical, respectively, which are commonly found in the fragmentation of Tyr containing peptide anions [34]. The fragment ion at m/z 71 indicates the loss of a neutral *p*-hydroxytoluene unit from the c_1 ion. Figure 4.4 compares the experimental IRMPD spectrum of the TyrPhe c$_1$ anion with theoretical spectra for several candidate structures. The 1500-1700 cm^{-1} region is critically diagnostic and it is immediately obvious that the best match is provided by structure **b**, a Tyr side-chain deprotonated isomer (analogous to structure VI in Figure 4.1). Structure **a** is isomerically identical to **b**, but conformationally different, and fails to match the experimental band near 1690 cm^{-1}. The difference in CO stretching frequency is due to different H-bonding in the two conformers: interaction with two NH$_2$ groups in **a** versus one in **b** induces a larger red shift of the C=O stretching mode. Relative energies of **a** and **b** are identical within the computational accuracy. One might argue that a small admixture of structures **c** (C-terminal

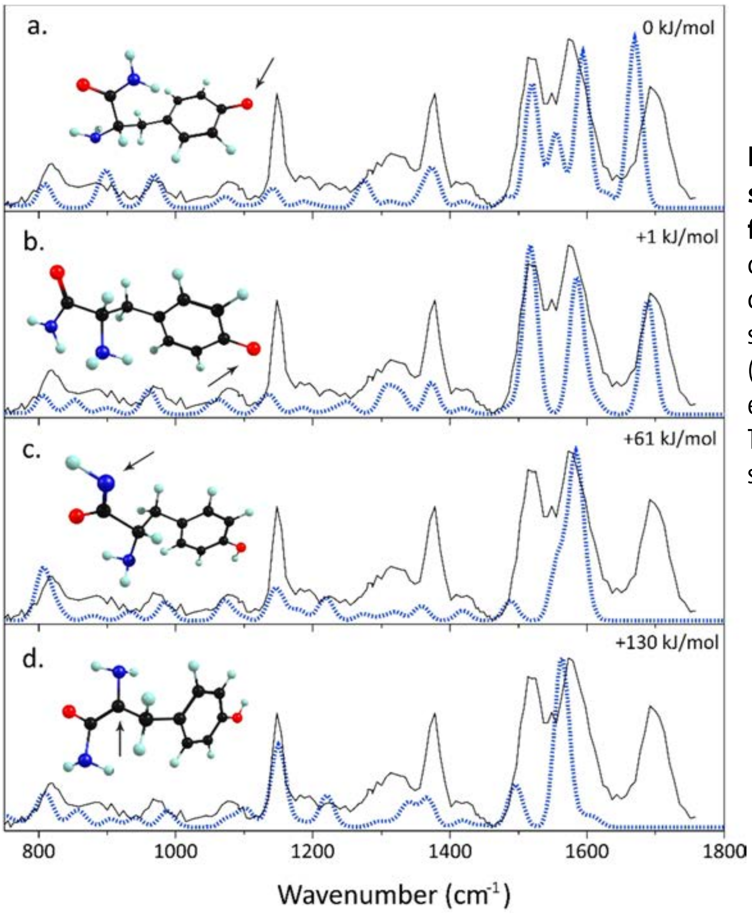


Figure 4.4. Experimental IRMPD spectrum of the c_1 fragment ion from [TyrPhe – H]$^-$ (black in all panels) compared with calculated spectra (dashed) for candidate structures deprotonated on the tyrosine side chain (**a** and **b**), the terminal amide nitrogen (**c**) and the α-carbon (**d**). Calculated relative free energy values are given for each of the structures. The arrow indicates the deprotonation site for each structure.

amidate) and/or **d** (deprotonated C_α enolate) could be present, although a higher intensity of the band at 1575 cm^{-1} would be expected if their abundance were substantial. Based purely on the computed relative free energies, their presence would be unlikely. The most intense absorptions in the computed spectrum for structure **b** are located at 1687, 1575 and 1519 cm^{-1}, corresponding to the carbonyl stretching mode and two different in-plane phenoxide C-O stretching modes coupled with aromatic CC stretch vibrations, respectively. The most intense absorptions in the computed spectrum for structure **b** are located at 1687, 1575 and 1519 cm^{-1}, corresponding to the carbonyl stretching mode and two different in-plane phenoxide C-O stretching modes coupled with aromatic CC stretch vibrations, Features at 1425 and 1378 cm^{-1} are also assigned to in-plane vibrations of the phenoxide moiety, and correspond favorably with similar vibrations in the phenoxide isomer of deprotonated *para*-hydroxybenzoic acid (*p*-HBA), for which an experimental gas-phase spectrum is available[62]. Lower-frequency bands at 1334 and 1313 cm^{-1} are due to coupled CC/NC stretching and C-H bending modes. The band at 1147 cm^{-1} stands out in the experimental spectrum but its intensity is poorly matched by the computed spectrum for structure **b**. According to the calculation, this band is due to the NH_2 twisting mode; such vibrational modes of an amino group have often been noticed to be poorly reproduced by harmonic frequency calculations[63], which is suspected to be due to the anharmonic nature of the potential for this vibration[64,65].

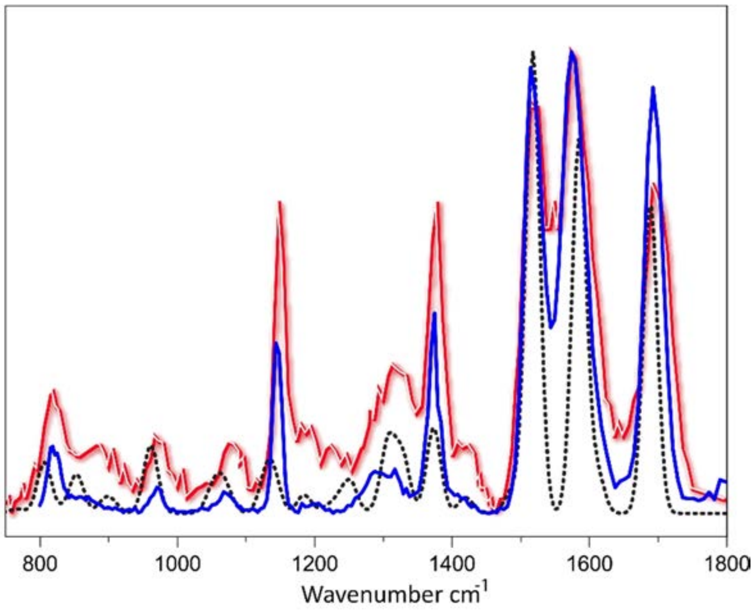


Figure 4.5. Comparison of the experimental IRMPD spectra of deprotonated Tyr-NH2 (blue, thick) and the c_1 anion from deprotonated TyrPhe (red, thin). The calculated spectrum for the side-chain deprotonated structure b from Figure 4.4 is shown as the dashed black line.

In order to further support our structure assignment of the TyrPhe c_1 anion, we recorded the IRMPD spectrum of the conjugate base of commercially available tyrosineamide (Tyr-NH_2). The results are shown in Figure 4.5. Clearly, the two experimental spectra are very similar, especially in band positions. In addition, IRMPD fragments for the c_1 anion from TyrPhe and for deprotonated Tyr-NH_2 are the same: m/z 106, 107 and 71. Hence, these results convincingly indicate that the c_1 fragment generated from TyrPhe has an amide C-terminal motif with deprotonation on the phenol side chain of the Tyr residue.

4.3.4. TrpPhe c_1-anion

Gas-phase acidities of the free amino acids Tyr and Trp are very similar around 1410 kJmol^{-1}, with Tyr being only slightly more acidic[66,67] (although the difference is smaller than the reported uncertainty). Perhaps more surprising, with a gas-phase acidity of around 1425 kJ mol^{-1}, phenol is only slightly (≈10 kJ mol^{-1}) more acidic than indole. Hence, deprotonation of the Trp side chain may be expected to compete with backbone deprotonation in *c*-type anions, similar to what is observed for the Tyr side chain.

Upon CO_2 laser irradiation, deprotonated TrpPhe was found to dissociate mainly into m/z 202, 164 and 147, corresponding to c_1, y_1 and z_1-type ions. IR irradiation of the isolated c_1 fragment anion subsequently produces m/z 130 and 71, ordered by their abundances. The m/z 130 fragment is assumed to correspond to a deprotonated methyl-substituted indole ring, while the m/z 71 fragment ion likely also results from side-chain detachment, but with the negative charge remaining on the backbone fragment.

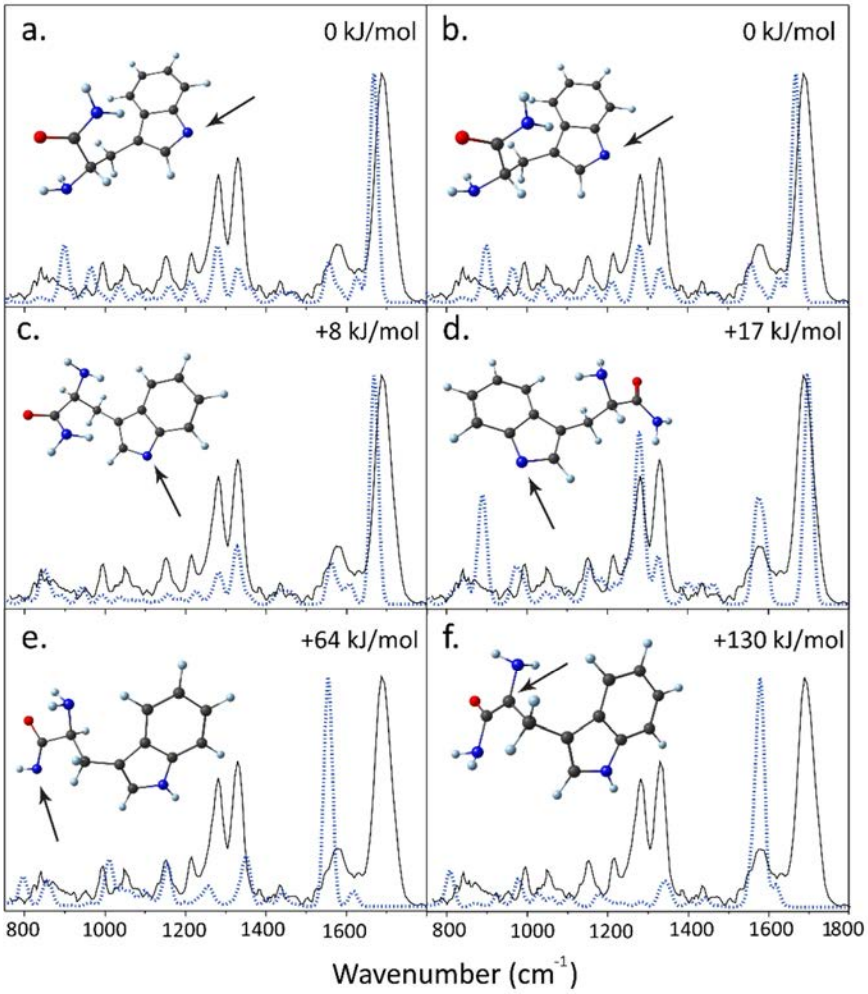


Figure 4.6. Experimental IRMPD spectrum of the c_1 fragment from deprotonated TrpPhe (black in all panels) compared with calculated spectra (dotted) for structures with a deprotonated indole ring (a, b, c, d), deprotonated C-terminal amide N-atom (e) and a deprotonated α-carbon atom (f). Calculated relative free energy values (in kJ/mol) are given in each panel. The arrow indicates the deprotonation site for each of the structures.

The IRMPD spectrum of this c_1-fragment anion compared to spectra calculated for candidate structures is presented in Figure 4.6. Structures **a**, **b**, **c** and **d** represent structures deprotonated on the indole nitrogen atom and they are only different in their conformations; their spectra differ mainly in relative band intensities. In structure **f** deprotonation occurs on the α-carbon atom and in **e** on the C-terminal amide nitrogen. From the comparisons in Figure 4.6 it is obvious that the theoretical spectra for structures **e** and **f** fail to match the experimental spectrum. In particular, both the amidate (**e**) and enolate (**f**) resonance structures cause a significant reduction of the carbonyl C=O bond order, red-shifting the strong CO stretching mode more than 100 cm^{-1} away from the experimental position. Furthermore, relative energies for the backbone deprotonated isomers are substantially higher than for the indole-deprotonated structures. Note that the relative energies are similar to the analogous isomers for the c_1 anion from TyrPhe, reflecting the comparable gas-phase acidities of phenol and indole.

Although all of the side-chain deprotonated conformers (**a**-**d**) in Figure 4.6 provide a reasonable, global match to the experimental spectrum, and none of them can be excluded as a possible contributor in a mixture of conformers, structure **c** is seen to match particularly well. The carbonyl stretching mode for this conformer appears at exactly the same position as for the c_1 anion from TyrPhe, 1687 cm^{-1}, indicating that both c_1 ions form very similar structures. Weak features at 1577 and 1565 cm^{-1} in the spectrum of conformer **c** are due to NH_2 scissoring vibrational modes from the C-terminal amide NH_2 group and from the N-terminal amine, respectively. Two bands with relatively high intensity in the experimental spectrum at 1281 and 1327 cm^{-1} match almost perfectly with theory and can be assigned as indole CN stretching coupled with aromatic CC stretching and aromatic in-plane CH bending vibrations, respectively. In conclusion, the IRMPD spectrum in combination with the calculations convincingly identifies this c_1 ion as a structure deprotonated on the tryptophan side chain.

4.3.5. c_2-anion from Leu-Ala^2-Enkephaline (YAGFL)

The c_1 ions studied above address the competition between deprotonation of the backbone C- terminal amide moiety and either the Tyr or Trp side chains. While the backbone in this case contains an amide group at the C-terminus, it does not posses an amide bond linking two residues. The study of the GlyGly c_2 anion indicates that deprotonation on the amide peptide linkage is more favorable than on the C-terminal amide group and that this motif is indeed formed in the dissociation reaction. Here we investigate the structure of a c_2 fragment containing a Tyr residue: the c_2 ion of deprotonated YAGFL.

Upon CO_2 laser irradiation for 0.5 seconds, deprotonated YAGFL dissociates mainly into b_1, b_2 and c_2 fragments. Upon subsequent IRMPD induced by the FEL, the isolated c_2 anion at m/z 250 dissociates into m/z 162 and 186, with the former fragment being much higher in abundance. This fragment likely represents the b_1-sequence ion, while the identity of the m/z 186 fragment is presently not clear; identification of the fragmentation products is however not required for obtaining an IR spectrum of the precursor.

Comparison of experimental and theoretical spectra in Figure 4.7 suggests that the c_2 ion adopts a C-terminal amide structure in which the Tyr side chain is deprotonated. Panels **a** and **b** show two different conformers for this phenoxide isomer, which are not only very close in energy, but also have nearly identical computed spectra. According to the calculated spectrum for structure **a**, the strongest absorption band observed at 1702 cm^{-1} corresponds to a convolution of two unresolved carbonyl stretching modes predicted at 1689 and 1698 cm^{-1}. Furthermore, the band near 1600 cm^{-1} is due to three vibrational modes calculated at 1591, 1557 and 1578 cm^{-1} due to NH_2 scissoring modes and a phenoxide CO/aromatic CC stretching mode. The feature observed near 1515 cm^{-1} is clearly also a convolution of more than one band, and the calculated spectra confirm this, although the calculated separation between the bands appears to be slightly off. These bands correspond to phenoxide CO stretching modes coupled to aromatic ring vibrations. Several weaker bands in the 1100 – 1400 cm^{-1} range are reproduced well by the calculations in panels **a** and **b**, providing further evidence for a deprotonated Tyr side chain structure for this c_2 anion. Some small discrepancies between experiment and theory are observed at

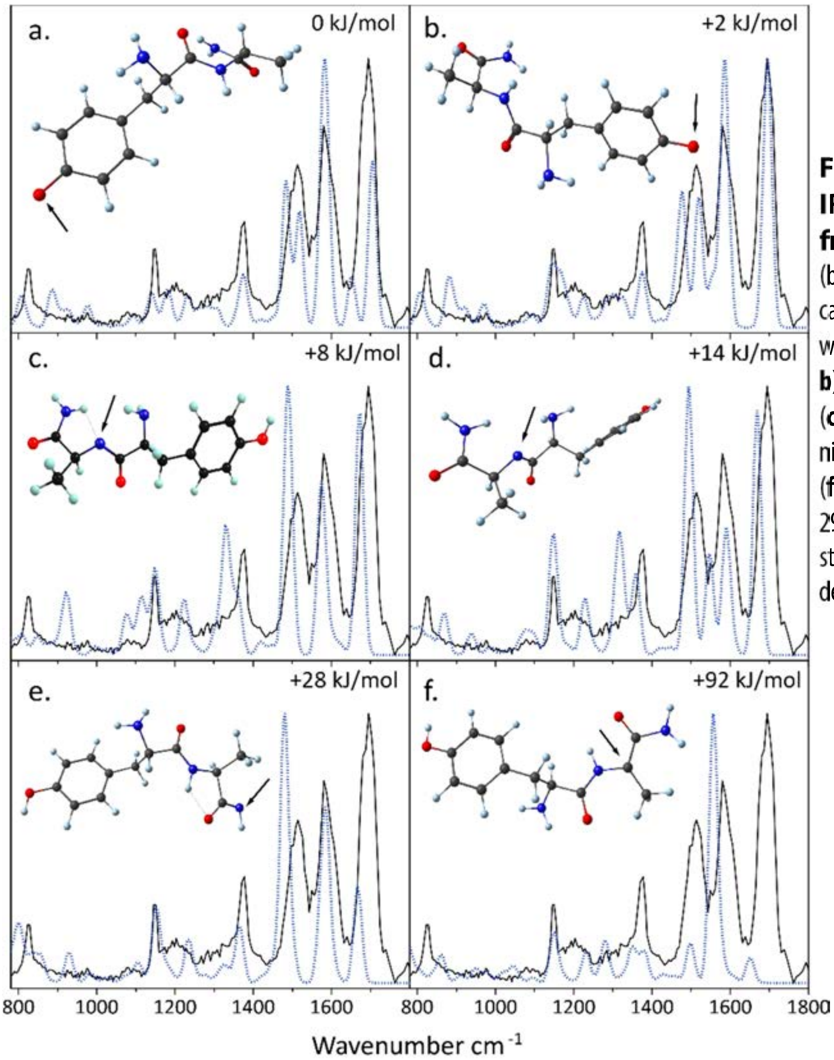


Figure 4.7. Experimental IRMPD spectrum of the c_2 fragment anion from YAGFL (black in all panels) compared with calculated spectra (dotted) for structures with a deprotonated tyrosine side chain (**a, b**), a deprotonated peptide bond nitrogen (**c, d**), a deprotonated C-terminal amide nitrogen (**e**) and a deprotonated α-carbon (**f**). Calculated relative free energy values at 298 K (in kJ mol^{-1}) are given for each of the structures. The arrow indicates the deprotonation site.

frequencies lower than 900 cm^{-1} which are likely due to inadequacies of the harmonic computations in predicting the vibrations of the NH_2 group [64,65].

Amidate structures in which the N-atom of the first amide group is deprotonated are only about 10 kJ/mol higher in energy. Two amidate conformers are shown in panels **c** and **d** of Figure 4.7 and their calculated band positions in the diagnostic 1450 – 1750 cm^{-1} range match only slightly worse than those for the phenoxide conformers of panels **a** and **b**. However, in the remainder of the spectrum, the two conformers of the amidate isomer feature strong calculated bands that are not observed in the experimental spectra, *e.g.* near 1320 and 910 cm^{-1}. Panels **e** and **f** show calculated spectra for a C-terminal amidate structure and an enolate structure, respectively. Their relative energies are substantially higher and the spectral mismatch with the experimental data is more apparent.

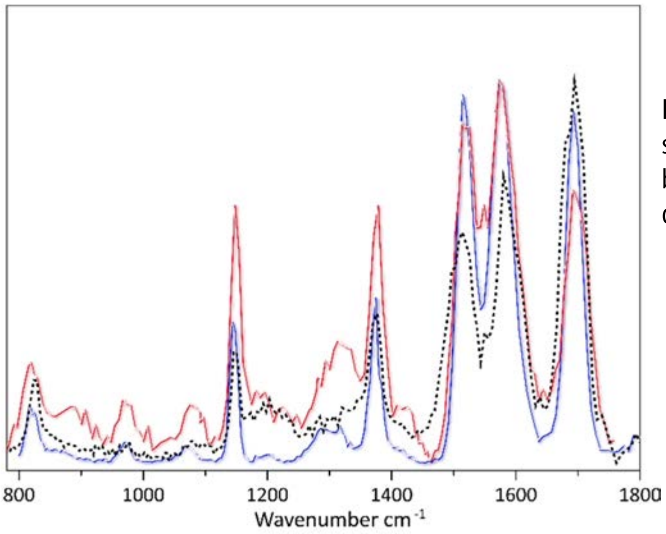


Figure 4.8: Comparison of the experimental IR spectra of the c_2 anion of deprotonated YAGFL (dotted black line), the c_1 anion of TyrPhe (red line) and deprotonated tyrosinamide (blue line).

Finally, further evidence for a phenoxide structure of this c_2 anion comes from overlaying its experimental spectrum with those of the c_1 fragment ion of TyrPhe and of deprotonated tyrosinamine ,Tyr-NH_2 (Figure 4.8). The six main bands in each of the three IR spectra match accurately in band position and bandwidth and even most of the minor bands agree. Deviations in relative intensities may sensitively depend on day-to-day variations in various experimental parameters, such as the laser power profile over the scan range.

4.4. Discussion

The IR spectra combined with the theoretical computations indicate that all *c*-ions studied correspond to truncated linear peptide structures with an amide –C(=O)-NH_2 moiety at the C-terminus, which becomes deprotonated only for c_1 anions without an acidic residue, such as observed here for the c_1 anion from AlaPhe. No evidence was found for the tautomeric but higher energy enol-imide motif –C(-OH)=NH that is often suggested as the *c*-ion structure resulting from ECD and ETD of protonated peptides. As the *c*-ions no longer contain a C-terminal carboxylic acid group, deprotonation has to occur elsewhere in the molecule. In previous MS^3 experiments on c_2 fragments containing only alkyl side chains, Harrison reported difficulties in distinguishing between deprotonation on the peptide bond amide nitrogen and the C-terminal amide nitrogen[5]. Here, the IR spectra show that for the c_2 ion from deprotonated GlyGlyPhe the thermodynamically most favored isomer is formed, *i.e.* an amidate structure where the amide nitrogen of the first peptide bond is deprotonated. No evidence was found for deprotonation on the C-terminal amide group.

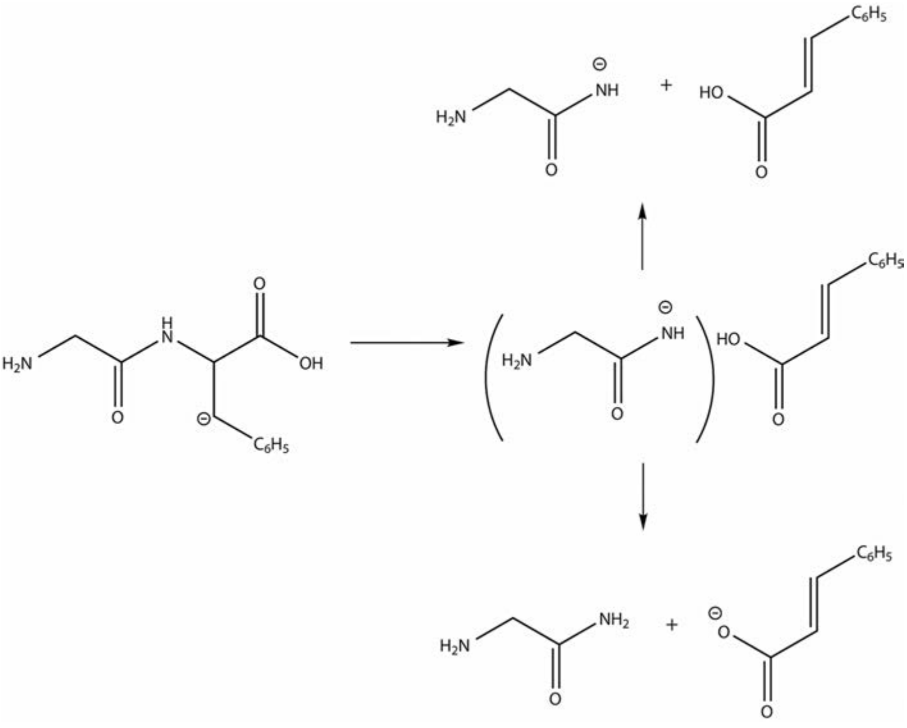


Figure 4.9. Mechanism for formation of c-type ions from anionic peptides containing Phe as the C-terminal residue, as proposed by Bowie.

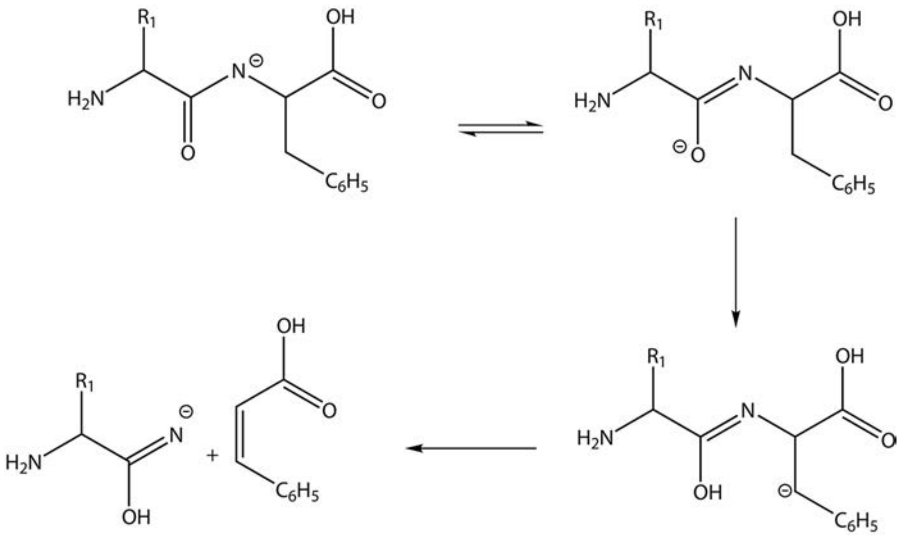


Figure 4.10. Mechanism for formation of c-type fragments proposed by Harrison for anionic peptides containing Phe as the C-terminal residue

Deprotonation at the peptide bond amidic nitrogen is indeed a common motif, as it was for instance identified as the structure of the deprotonated methyl ester of tri-alalnine[29] and of *a*-type anions of poly-alanine[41,61].

For c_1 ions without an acidic side chain, deprotonation of the C-terminal amide group appears to be the only possible structure. Harrison reported abundant c_1 ion formation from deprotonated GlyPhe at relatively high activation energies and proposed deprotonation at the C-terminal amide nitrogen, though in the tautomeric enol-imide form[5]. Note that Bowie and coworkers actually suggested the alternative, lower energy keto-tautomer for this c_1 ion[34]. The proposed mechanisms of these authors are reproduced in Figures 4.9. and 4.10. In the present experiments, the c_1 fragment ion from AlaPhe is shown to contain a deprotonated C-terminal amide moiety in its keto form. On the other hand, c_1 ions with a Tyr or Trp side chain from the Phe containing dipeptide were actually observed in our instrument and their IR spectra indicate that they are deprotonated on the phenol and indole side chains, respectively. This suggests either that an alternative

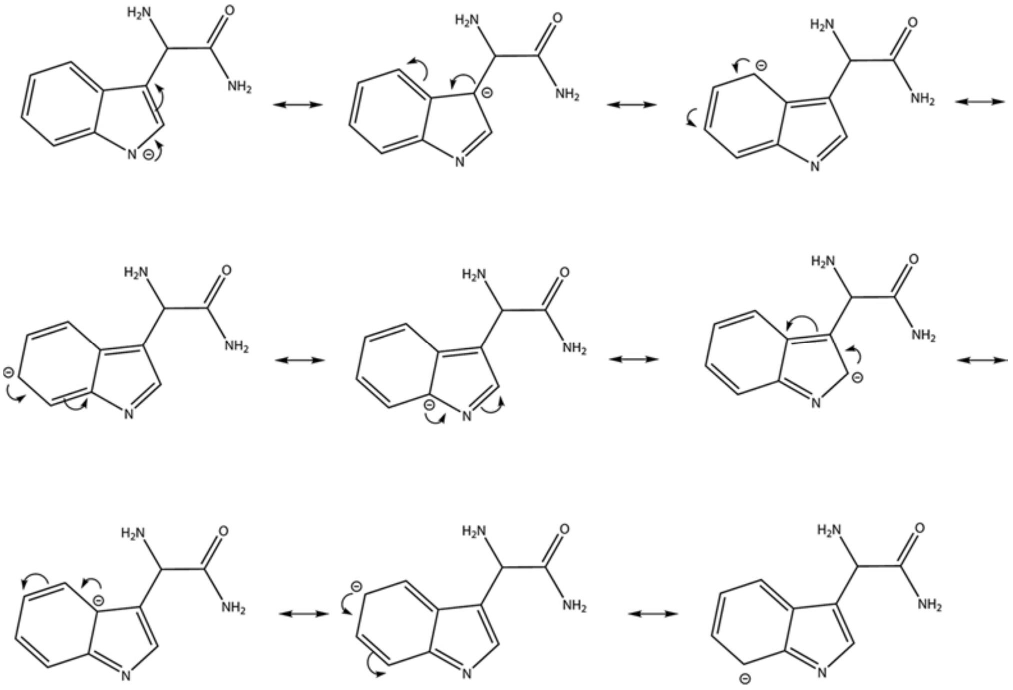


Figure 4.11. Illustration of charge delocalization in the c_1 fragment from deprotonated TrpPhe.

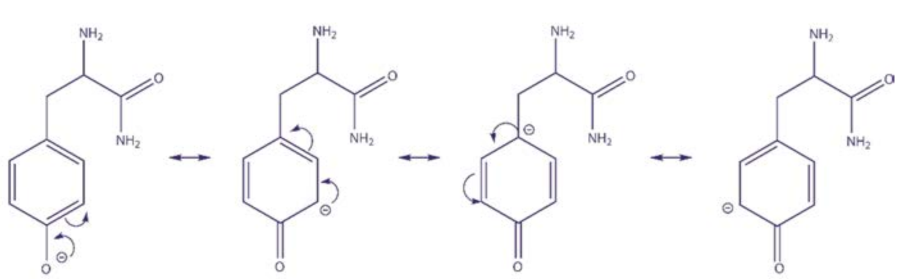


Figure 4.12. Illustration of charge delocalization in the c_1 fragment from deprotonated TyrPhe.

mechanism than that of Harrison and/or Bowie is at play for these peptides (*e.g.* a charge-remote mechanism), or that isomerization to the lower energy form occurs after the Bowie/Harrison processes displayed in Figures 4.9. and 4.10.

For the c_2 ion from deprotonated YAGFL a similar phenol-deprotonated structure was found. Thermodynamically, this can be understood from the substantially higher gas-phase acidity of phenol ($\Delta G \approx 1429$ kJ/mol [68]) as compared to that of N-methylacetamide ($\Delta G \approx 1483$ kJ/mol) serving as a model for the acidity of the peptide amide bond proton. Based on these arguments and the gas-phase acidity of indole ($\Delta G \approx 1435$ kJ/mol), anionic *c*-type fragments containing tryptophan are then also predicted to deprotonate on the side chain rather than on the backbone. The high gas-phase acidity of phenol and indole is the result of the efficient delocalization of the negative charge over the entire phenol and indole systems[69], as illustrated in Figures 4.11. and 4.12. We note finally that in contrast, for anionic *b*-type peptide fragments, deprotonation of the oxazolone ring was found to dominate over deprotonation at the Tyr side chain, as also established using infrared ion spectroscopy[42].

4.5. Conclusion

IRMPD spectroscopy using the FELIX free electron laser was applied to investigate the structure of five short anionic *c*-type fragments. By comparison with computed IR spectra for different candidate isomers, the structures were found to be truncated peptides with an amide moiety (in the normal keto-amino form) at the C-terminus. Deprotonation occurs at different sites depending on the residues involved. For c_2 ions without acidic residues, the amide linkage between the two residues is deprotonated at the amidic nitrogen. However, incorporation of mildly acidic residues Tyr and Trp results in side-chain deprotonated structures. Deprotonation of the C-terminal amide is observed only in the case of c_1 ions not containing acidic residues.

As a future outlook, we note that it will be of interest to investigate the structure of larger *c*-ions from deprotonated peptides. In experiments on longer *c*-type fragments by Harrison, cyclization was not suggested to occur. This is an interesting characteristic of these *c*-type ions, as it may avoid possible problems related to sequence scrambling in sequencing applications[6,7].

Acknowledgment

We gratefully thank the skillful support by the FELIX staff, in particular by Drs. B. Redlich, A.F.G. van der Meer, and G. Berden. This work is part of the research program of FOM, which is financially supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). Additional support by NWO in the form of VI grant number 724.011.002 is gratefully acknowledged. JO thanks the Stichting Physica for support.

References

(1) McLafferty, F. W. *Tandem mass spectrometry (MS/MS): a promising new analytical technique for specific component determination in complex mixtures* Acc. Chem. Res. **1980**, *13*, 33.
(2) Siuzdak, G. *The emergence of mass-spectrometry in biochemical research* Proc. Natl. Acad. Sci. USA **1994**, *91*, 11290.
(3) Yates, J. R. *Mass spectrometry and the age of the proteome* J. Mass Spectrom. **1998**, *33*, 1.
(4) Biemann, K. *Contributions of mass spectrometry to peptide and protein structure* Biol. Mass Spectrom. **1988**, *16*, 99.
(5) Harrison, A. *Effect of phenylalanine on the fragmentation of deprotonated peptides* J. Am. Soc. Mass. Spectrom. **2002**, *13*, 1242.
(6) Harrison, A. G. *Peptide Sequence Scrambling Through Cyclization of* b_5 *Ions* J. Am. Soc. Mass. Spectrom. **2008**, *19*, 1776.
(7) Bleiholder, C.; Osburn, S.; Williams, T. D.; Suhai, S. n.; Van Stipdonk, M.; Harrison, A. G.; Paizs, B. l. *Sequence-Scrambling Fragmentation Pathways of Protonated Peptides* J. Am. Chem. Soc. **2008**, *130*, 17774.
(8) Harrison, A.; Young, A. *Fragmentation of deprotonated N-benzoylpeptides: Formation of deprotonated oxazolones* J. Am. Soc. Mass. Spectrom. **2004**, *15*, 446.
(9) Chass, G. A.; Marai, C. N. J.; Setiadi, D. H.; Csizmadia, I. G.; Harrison, A. G. *A Hartree-Fock, MP2 and DFT computational study of the structures and energies of "*b_2 *ions derived from deprotonated peptides. A comparison of method and basis set used on relative product stabilities* J. Mol. Struct. **2004**, *675*, 149.
(10) Harrison, A. G.; Young, A. B. *Fragmentation reactions of deprotonated peptides containing aspartic acid* Int. J. Mass Spectrom. **2006**, *255-256*, 111.
(11) Barlow, C. K.; O'Hair, R. A. J. *Gas-Phase Peptide Fragmentation: How Understanding the Fundamentals Provides a Springboard to Develping New Chemistry and Novel Proteomics Tools* J. Mass Spectrom. **2008**, *43*, 1301.
(12) Polfer, N. C.; Oomens, J. *Vibrational spectroscopy of bare and solvated ionic complexes of biological relevance* Mass Spectrom. Rev. **2009**, *28*, 468.
(13) Polfer, N. C.; Oomens, J. *Reaction products in mass spectrometry elucidated with infrared spectroscopy* Phys. Chem. Chem. Phys. **2007**, *9*, 3804.
(14) Eyler, J. R. *Infrared multiple photon dissociation spectroscopy of ions in Penning traps* Mass Spectrom. Rev. **2009**, *28*, 448.
(15) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *Spectroscopic and Theoretical Evidence for Oxazolone Ring Formation in Collision-Induced Dissociation of Peptides* J. Am. Chem. Soc. **2005**, *127*, 17154.
(16) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *Infrared spectroscopy and theoretical studies on gas-phase protonated Leu-enkephalin and its fragments: Direct experimental evidence for the mobile proton* J. Am. Chem. Soc. **2007**, *129*, 5887.
(17) Paizs, B.; Suhai, S. *Fragmentation pathways of protonated peptides* Mass Spectrom. Rev. **2005**, *24*, 508.
(18) Armentrout, P. B.; heaton, A. L. *Thermodynamics and Mechanisms of Protonated Diglycine Decomposition: A Guided Ion Beam Study* J. Am. Soc. Mass Spectrom. **2012**, *in press*.
(19) Oomens, J.; Young, S.; Molesworth, S.; van Stipdonk, M. *Spectroscopic Evidence for an Oxazolone Structure of the* b_2 *Fragment Ion from Protonated Tri-Alanine* J. Am. Soc. Mass. Spectrom. **2009**, *20*, 334.
(20) Yoon, S. H.; Chamot-Rooke, J.; Perkins, B. R.; Hilderbrand, A. E.; Poutsma, J. C.; Wysocki, V. H. *IRMPD Spectroscopy Shows That AGG Forms an Oxazolone* b_2^+ *Ion* J. Am. Chem. Soc. **2008**, *130*, 17644.

(21) Perkins, B. R.; Chamot-Rooke, J.; Yoon, S. H.; Gucinski, A. C.; Somogyi, A. r. d.; Wysocki, V. H. *Evidence of Diketopiperazine and Oxazolone Structures for HA b_2^+ Ion* J. Am. Chem. Soc. **2009**, *131*, 17528.

(22) Molesworth, S.; Osburn, S.; Van Stipdonk, M. *Influence of Size on Apparent Scrambling of Sequence During CID of b-Type Ions* J. Am. Soc. Mass. Spectrom. **2009**, *20*, 2174.

(23) Chen, X.; Yu, L.; Steill, J. D.; Oomens, J.; Polfer, N. C. *Effect of Peptide Fragment Size on the Propensity of Cyclization in Collision-Induced Dissociation: Oligoglycine b_2-b_8* J. Am. Chem. Soc. **2009**, *131*, 18272.

(24) Erlekam, U.; Bythell, B. J.; Scuderi, D.; Van Stipdonk, M.; Paizs, B.; Maitre, P. *Infrared Spectroscopy of Fragments of Protonated Peptides: Direct Evidence for Macrocyclic Structures of b_5 Ions* J. Am. Chem. Soc. **2009**, *131*, 11503.

(25) Verkerk, U. H.; Siu, C.-K.; Steill, J. D.; El Aribi, H.; Zhao, J.; Rodriquez, C. F.; Oomens, J.; Hopkinson, A. C.; Siu, K. W. M. *a_2 Ion Derived from Triglycine: An N1-Protonated 4-Imidazolidinone* J. Phys. Chem. Lett. **2010**, *1*, 868.

(26) Bythell, B. J.; Maitre, P.; Paizs, B. *Cyclization and Rearrangement Reactions of an Fragment Ions of Protonated Peptides* J. Am. Chem. Soc. **2010**, *132*, 14766.

(27) Andreazza, H. J.; Bowie, J. H. *The application of negative ion electrospray mass spectrometry for the sequencing of underivatized disulfide-containing proteins: insulin and lysozyme* Phys. Chem. Chem. Phys. **2010**, *12*, 13400.

(28) Brinkworth, C.; Bowie, J.; Bilusich, D.; Tyler, M. *The rothein peptides from the skin secretion of Roth's tree frog Litoria rothii. Sequence determination using positive and negative ion electrospray mass spectrometry* Rapid Comm. Mass Spectrom. **2005**, *19*, 2716.

(29) Harrison, A. G. *Sequence-specific fragmentation of deprotonated peptides containing H or alkyl side chains* J. Am. Soc. Mass. Spectrom. **2001**, *12*, 1.

(30) Li, Z.; Yalcin, T.; Cassady, C. J. *C-terminal amino acid residue loss for deprotonated peptide ions containing glutamic acid, aspartic acid, or serine residues at the C-terminus* J. Mass Spectrom. **2006**, *41*, 939.

(31) Men, L.; Wang, Y. *Fragmentation of the deprotonated ions of peptides containing cysteine, cysteine sulfinic acid, cysteine sulfonic acid, aspartic acid, and glutamic acid* Rapid Commun. Mass Spectrom. **2006**, *20*, 777.

(32) Pu, D.; Cassady, C. J. *Negative ion dissociation of peptides containing hydroxyl side chains* Rapid Commun. Mass Spectrom. **2008**, *22*, 91.

(33) Pu, D.; Clipston, N. L.; Cassady, C. J. *A comparison of positive and negative ion collision-induced dissociation for model heptapeptides with one basic residue* J. Mass Spectrom. **2010**, *45*, 297.

(34) Waugh, R. J.; Bowie, J. H.; Hayes, R. N. *Collision induced dissociations of deprotonated peptides: Dipeptides containing phenylalanine, tyrosine, histidine and tryptophan* Int. J. Mass Spectrom. **1991**, *107*, 333.

(35) Brinkworth, C. S.; Dua, S.; McAnoy, A. M.; Bowie, J. H. *Negative ion fragmentations of deprotonated peptides: backbone cleavages directed through both Asp and Glu* Rapid Commun. Mass Spectrom. **2001**, *15*, 1965.

(36) Harrison, A. G.; Siu, K. W. M.; El Aribi, H. *Amide bond cleavage in deprotonated tripeptides: a newly discovered pathway to "b_2 ions* Rapid Commun. Mass Spectrom. **2003**, *17*, 869.

(37) Yury, V. V.; Figard, B. J.; Morré, J.; Deinzer, M. L. *Fragmentation of peptide negative molecular ions induced by resonance electron capture* J. Chem. Phys. **2009**, *131*, 044317.

(38) Waugh, R. J.; Bowie, J. H.; Gross, M. L. *Collision-Induced Dissociations of Deprotonated Peptides. Dipeptides Containing Asn, Arg and Lys* Australian Journal of Chemistry **1993**, *46*, 693.

(39) Andreazza, H. J.; Wang, T.; Bilusich, D.; Hoffmann, P.; Bowie, J. H. *Negative ion fragmentations of deprotonated peptides containing post-translational modifications: diphosphorylated systems containing Ser, Thr and Tyr. A characteristic phosphate/phosphate cyclisation. A joint experimental and theoretical study* Rapid Commun. Mass Spectrom. **2009**, *23*, 1825.

(40) Yoo, H.; Wang, N.; Zhuang, S.; Song, H.; Hakansson, K. *Negative-Ion Electron Capture Dissociation: Radical-Driven Fragmentation of Charge-Increased Gaseous Peptide Anions* J. Am. Chem. Soc. **2011**, *133*, 16790.

(41) Oomens, J.; Steill, J. D. *The Structure of Deprotonated Tri-Alanine and Its a_3^- Fragment Anion by IR Spectroscopy* J. Am. Soc. Mass. Spectrom. **2010**, *21*, 698.

(42) Grzetic, J.; Oomens, J. *Spectroscopic evidence for an oxazolone structure in anionic b-type peptide fragments* J. Am. Soc. Mass Spectrom. **2012**, *23*, 290.

(43) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process* J. Am. Chem. Soc. **1998**, *120*, 3265.

(44) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations* Anal. Chem. **2000**, *72*, 563.

(45) Cooper, H. J.; Håkansson, K.; Marshall, A. G. *The role of electron capture dissociation in biomolecular analysis* Mass Spectrom. Rev. **2005**, *24*, 201.

(46) Coon, J. J. *Collisions or electrons? Protein sequence analysis in the 21st century* Anal. Chem. **2009**, *81*, 3208.

(47) Zubarev, R. A. *Reactions of polypeptide ions with electrons in the gas phase* Mass Spectrom. Rev. **2003**, *22*, 57.

(48) Brueker, K.; Oh, H. B.; Lin, C.; Carpenter, B. K.; McLafferty, F. W. *Nonergodic and conformational control of the electron capture dissociation of protein cations* Proc. Nat. Acad. Sci. USA **2004**, *101*, 14011.

(49) Syrstad, E. A.; Turecek, F. *Toward a General Mechanism of Electron Capture Dissociation* J. Am. Soc. Mass Spectrom. **2005**, *16*, 208.

(50) Skurski, P.; Sobcyk, M.; Jakowski, J.; Simons, J. *Possible mechanisms for protecting N-Cα bonds in helical peptides from electron-capture (or transfer) dissociation* Int. J. Mass Spectrom. **2007**, *265*, 197.

(51) Frison, G.; van der Rest, G.; Turecek, F.; Besson, T.; Lemaire, J.; Maitre, P.; Chamot-Rooke, J. *Structure of Electron-Capture Dissociation Fragments from Charge-Tagged Peptides Probed by Tunable Infrared Multiple Photon Dissociation* J. Am. Chem. Soc. **2008**, *130*, 14916.

(52) Valle, J. J.; Eyler, J. R.; Oomens, J.; Moore, D. T.; van der Meer, A. F. G.; von Helden, G.; Meijer, G.; Hendrickson, C. L.; Marshall, A. G.; Blakney, G. T. *Free electron laser-Fourier transform ion cyclotron resonance mass spectrometry facility for obtaining infrared multiphoton dissociation spectra of gaseous ions* Rev. Sci. Instrum. **2005**, *76*, 023103.

(53) Marshall, A. G.; Wang, T. C. L.; Ricca, T. L. *Tailored excitation for Fourier transform ion cyclotron mass spectrometry* J. Am. Chem. Soc. **1985**, *107*, 7893.

(54) Oepts, D.; van der Meer, A. F. G.; van Amersfoort, P. W. *The infrared free-electron laser facility FELIX* Infrared Phys. Technol. **1995**, *36*, 297.

(55) Oomens, J.; Sartakov, B. G.; Meijer, G.; von Helden, G. *Gas-phase infrared multiple photon dissociation spectroscopy of mass-selected molecular ions* Int. J. Mass Spectrom. **2006**, *254*, 1.

(56) Polfer, N. C. *Infrared multiple photon dissociation spectroscopy of trapped ions* Chem. Soc. Rev. **2011**, *40*, 2211.

(57) MacAleese, L.; Maitre, P. *Infrared Spectroscopy of Organometallic Ions in the Gas Phase* Mass Spectrom. Rev. **2007**, *26*, 583.

(58) Mize, T. H.; Taban, I.; Duursma, M.; Seynen, M.; Konijnenburg, M.; Vijftigschild, A.; Doornik, C. V.; Rooij, G. V.; Heeren, R. M. A. *A modular data and control system to improve sensitivity, selectivity, speed of analysis, ease of use, and transient duration in an external source FTICR-MS* Int. J. Mass Spectrom. **2004**, *235*, 243.

(59) Merrill, G. N.; Kass, S. R. *Calculated Gas-Phase Acidities Using Density Functional Theory: Is It Reliable?* J. Phys. Chem. **1996**, *100*, 17465.

(60) Oomens, J.; Steill, J. D. *Free carboxylate stretching modes* J. Phys. Chem. A **2008**, *112*, 3281.

(61) Chass, G. A.; Marai, C. N. J.; Harrison, A. G.; Csizmadia, I. G. *Fragmentation reactions of a_2 ions derived from deprotonated dipeptides - A synergy between experiment and theory* J. Phys. Chem. A **2002**, *106*, 9695.

(62) Steill, J. D.; Oomens, J. *Gas-Phase Deprotonation of p-Hydroxybenzoic Acid Investigated by IR Spectroscopy: Solution-Phase Structure Is Retained upon ESI* J. Am. Chem. Soc. 2009, *131* (38), 13570–13571
2009, *131*, 13570.

(63) Oomens, J.; Steill, J. D.; Redlich, B. *Gas-phase IR spectroscopy of deprotonated amino acids* J. Am. Chem. Soc. **2009**, *131*, 4310.

(64) Sinclair, W. E.; Pratt, D. W. *Structure and vibrational dynamics of aniline and aniline–Ar from high resolution electronic spectroscopy in the gas phase* J. Chem. Phys. **1996**, *105*, 7942.

(65) Piest, H.; von Helden, G.; Meijer, G. *Infrared spectroscopy of jet-cooled neutral and ionized aniline–Ar* J. Chem. Phys. **1999**, *110*, 2010.

(66) O'Hair, R. A. J.; Bowie, J. H.; Gronert, S. *Gas phase acidities of the α amino acids* Int. J. Mass Spectrom. Ion Process. **1992**, *117*, 23.

(67) Jones, C. M.; Bernier, M.; Carson, E.; Colyer, K. E.; Metz, R.; Pawlow, A.; Wischow, E. D.; Webb, I.; Andriole, E. J.; Poutsma, J. C. *Gas-phase acidities of the 20 protein amino acids* Int. J. Mass Spectrom. **2007**, *267*, 54.

(68) Gas-phase acidities are average values from the NIST Chemistry Webbook.

(69) McMahon, T. B.; Kebarle, P. *Intrinsic acidities of substituted phenols and benzoic acids dtermined by gas-phase proton-transfer equilibria* J. Am. Chem. Soc. **1977**, *99*, 2222.

Chapter 5.

Spectroscopic evidence for an oxazolone structure in anionic *b*-type peptide fragments

Infrared spectra of anionic *b*-type fragments generated by collision induced dissociation (CID) from deprotonated peptides are reported. Spectra of the b_2 fragments of deprotonated AlaAlaAla and AlaTyrAla have been recorded over the 800-1800 cm^{-1} spectral range by multiple-photon dissociation (MPD) spectroscopy using an FTICR mass spectrometer in combination with the free electron laser FELIX. Structural characterization of the *b*-type fragments is accomplished by comparison with density functional theory calculated spectra at the B3LYP/6-31++G(d,p) level for different isomeric structures. Although diketopiperazine structures represent the energetically lowest isomers, the IR spectra suggest an oxazolone structure for the b_2 fragments of both peptides. Deprotonation is shown to occur on the oxazolone α-carbon, which leads to a conjugated structure in which the negative charge is practically delocalized over the entire oxazolone ring, providing enhanced gas-phase stability.

*Adapted from: J.Grzetic and J.Oomens, *Spectroscopic evidence for an oxazolone structure in anionic b-type peptide fragments*, J. Am. Soc. Mass Specctrom.,(2012), 23, 290-300

5.1. Introduction

Collision induced dissociation (CID) tandem mass spectrometry has become an essential tool in proteomics as it is routinely used to determine peptide and protein sequences. In parallel to these application-oriented developments, the reaction chemistry underlying peptide fragmentation has been a topic of much debate over the past decades[1-4]. A good example of where such detailed mechanistic information may aid in the practical analysis of peptide tandem mass spectra are the complications arising from possible sequence scrambling; CID fragments from protonated peptides may adopt cyclic structures in the gas phase, which can reopen at various positions resulting in a permutation of the initial peptide sequence[5-7].

An obviously important parameter in understanding the fragmentation chemistry is knowledge of the molecular structures of the fragment ions formed in the CID process, which can be ambiguous based on the m/z information alone. The combination of tunable infrared lasers with tandem mass spectrometers has made it possible to record IR spectra of CID peptide fragments, which can be related to their structure generally with high confidence. Such detailed structural information in combination with results from computational chemistry can then be used to establish the CID reaction pathways.

Spectroscopic structure determination has thus far been applied predominantly to *b*-type CID fragment ions generated from protonated peptides. Various different structures have been identified depending on the properties of their amino acid constituents, peptide length, as well as on the collision energy. The b_4 fragment ion from protonated Leucine-Enkephaline was shown to possess an oxazolone structure, which can be protonated at different positions[8,9]. While b_2 fragments had been suggested to form diketopiperazine structures[10], IR spectroscopy definitively determined an oxazolone structure for the b_2 fragments from protonated Ala_3 (AAA) and AlaGlyGly (AGG)[11,12] and a diketopiperazine structure for the b_2 ion from $HisAla_4$[13]. For larger *b*-ions \`head-to-tail' cyclization produces macrocyclic structures which may result in sequence scrambling[5,6,14,15]; recent IR spectroscopic studies have indeed revealed the abundant occurrence of such structures[16-18]. Still more recently, *a*-type fragment ions from protonated peptides were spectroscopically shown to form cyclic structures as well[19,20]. This behavior of *a*-type fragments may also lead to sequence scrambling in MS^n peptide analysis.

Although most peptide sequencing applications make use of protonated peptide species, current MS instrumentation can generally run in both positive and negative ion mode. Hence, sequence coverage could readily be increased by analyzing the same peptide in negative ion mode and using the complementary information. Moreover, for peptides containing many acidic residues or for phosphorylated peptides, running in negative ion mode, i.e. sequencing the deprotonated peptide, may even be more efficient.

The first IR spectroscopic investigation of deprotonated peptide fragments considered the a_3 fragment from Ala_3, and the main issue was to determine the deprotonation site in the N-terminal fragment, which no longer possesses a carboxylate moiety at the C-terminus. Resolving the deprotonation site is important as dissociation reactions are generally charge-directed. The IR spectra showed unambiguously that the a_3^- fragment has a deprotonated amide (amidate) structure[21] as proposed earlier by Harrison and coworkers[22].

As for protonated peptides, amide bond cleavage is a common peptide fragmentation pattern for deprotonated peptides, leading to the formation of *b*-type ions. What are the structures of these *b*-type fragments from deprotonated peptides? Do they also form oxazolone or diketopiperazine structures as is now well-known for protonated peptides? Or are there other structures that are more likely in the case of deprotonated peptides? Moreover, where do these N-terminal fragments deprotonate in the absence of acidic residues? These are the questions that we will address here by comparing experimental IRMPD spectra for deprotonated b_2 fragment ions with spectra computed for candidate structures using density functional theory (DFT).

Formation of *b*-type fragments from deprotonated peptides was first reported by Heerma and co-workers[23,24] and there are several proposed mechanisms for their formation, which all rely on proton transfer in the first stage of the reaction. Although this "mobile proton theory" may appear counterintuitive for a deprotonated peptide – having one proton fewer, not more – Bache and coworkers showed explicitly that hydrogen atom scrambling also occurs in anionic peptides upon collisional activation[25].

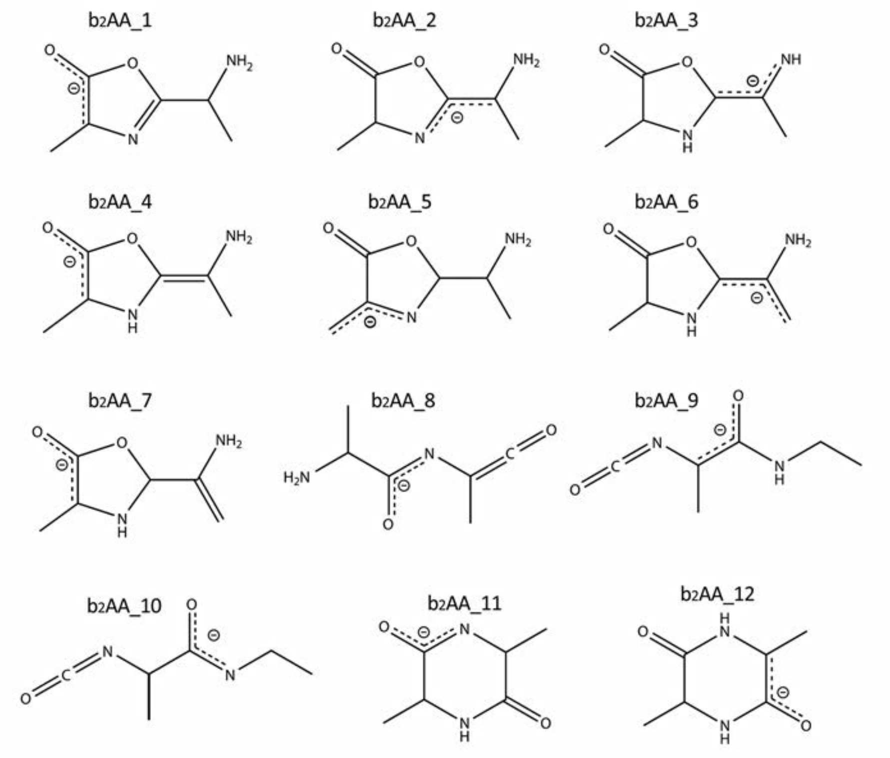
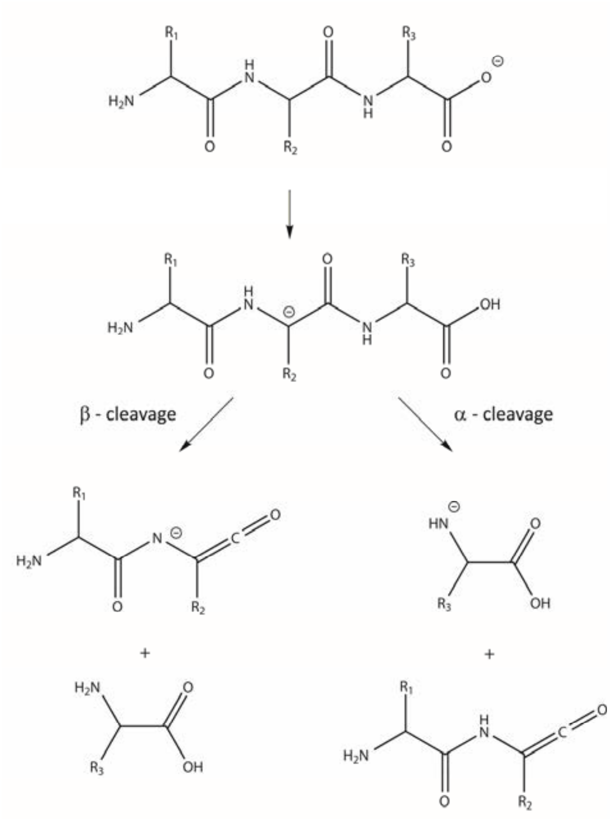
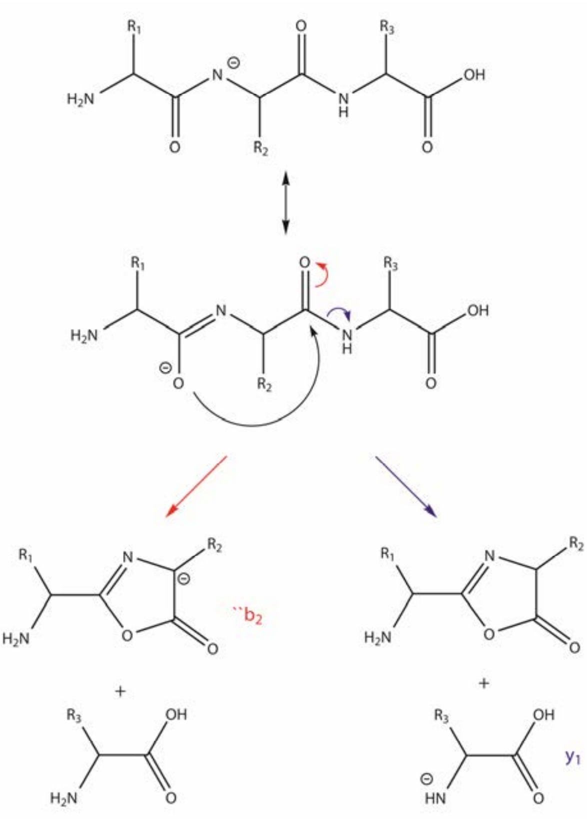


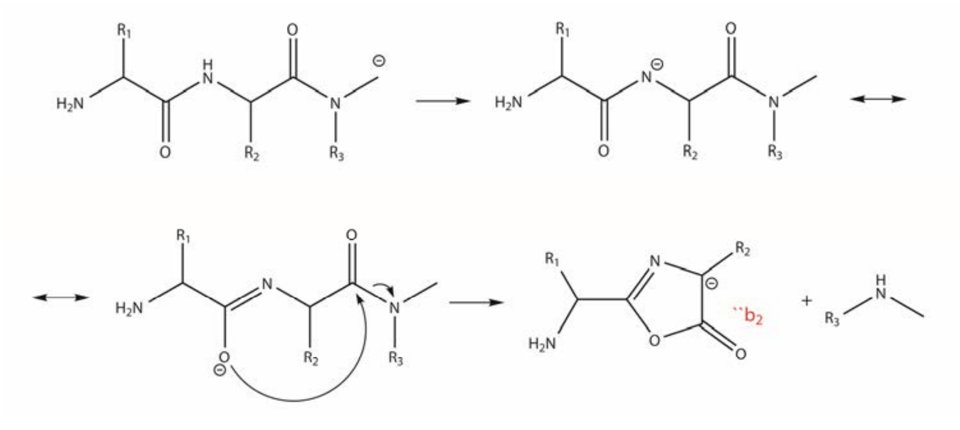
Figure 5.1. Suggested structures for the b_2 ion.



Scheme 5.1. Proposed mechanism for fragmentation of deprotonated peptides into *b* and *y* fragments as suggested by Bowie[27].



Scheme 5.2. Mechanism for formation of b_2 oxazolone fragments from deprotonated peptides as proposed by Harrison and coworkers.



Scheme 5.3. Alternative mechanism to form a b_2 oxazolone structure, involving the a_3^- fragment as an intermediate, as proposed particularly for residues without an amidic hydrogen by Harrison et al.[29].

Reaction pathways involving amide bond cleavage as proposed by Bowie and co-workers are depicted in Scheme 5.1 (for the formation of the b_2 ion) and result in a linear amidate *b*-ion with a C=C=O unit at the C-terminus[26-28]. According to molecular modeling studies at the AM1 level of theory, backbone cleavage leading to *b*-type ions occurs via an enolate intermediate, which is produced by proton transfer to the COO^- moiety[28]. In this mechanism, so called α-cleavage of the peptide backbone leads to an N-deprotonated amino acid (y_1 ion), and β-cleavage forms the charged N-terminal fragment (b_2 ion).

An alternative pathway proposed by Harrison and co-workers, depicted in Scheme 5.2, results in a b_2 fragment with a cyclic oxazolone structure[29-32]. Harrison actually proposes two alternative mechanisms. In the first one, the *b*-ion is formed directly from the parent deprotonated peptide and in the second one, the b_2-ion is a sequential fragment formed via the a_3 fragment. In Scheme 5.2, the precursor ion is the deprotonated peptide and an amidate structure is formed by proton transfer to the COO^- moiety. This proton transfer is followed by nucleophilic attack of the conjugated negatively charged oxygen atom onto a carbonyl carbon, forming an oxazolone ring, which can retain the negative charge. The same pathway can lead to the y_1 fragment if the negative charge remains on the C-terminal fragment. Formation of b_2 ions via the a_3-fragment also occurs via an amidate intermediate structure[32], see Scheme 5.3, although this pathway was suggested to be active especially for residues without an amidic hydrogen, such as Pro[29]. The b_2 fragment can then be formed via a similar nucleophilic attack.

To gain further insight into the possible structures of b_2 ions, Harrison and coworkers turned to computational chemistry[33]. In this study, the energetically most favorable b_2 ion structures (for the GG sequence) correspond to the structures b_2AA_1, b_2AA_8, b_2AA_9, b_2AA_11 and b_2AA_12 presented in Figure 5.1. Inspired by Harrison's findings, we here report on the results of an IR spectroscopic investigation of the b_2 ion derived from the deprotonated peptides tri-alanine (AAA) and alanyl-tyrosil-alanine (AYA).

5.2. Experimental and computational methods

All experiments were conducted using the free electron laser FELIX[34] and a home-built FTICR mass spectrometer equipped with a 4.7 T actively-shielded superconducting magnet (Cryomagnetics Inc., Oak Ridge, TN, USA)[35-37].

Deprotonated peptide anions were generated by electrospray ionization (Micromass Z-Spray source) using solutions containing about 1 mM of the peptide and 1 mM NaOH to enhance the formation of peptide anions. Tri-alanine and alanyl-tyrosyl-alanine were obtained from Sigma-Aldrich (Zwijndrecht, The Netherlands) and Bachem (Weil am Rhein, Germany), respectively. A mixture of methanol and water (7/3 ratio) was used as solvent. CID fragments were generated by nozzle-skimmer dissociation in the high-pressure region at the inlet of the mass spectrometer. Ions were accumulated in a linear hexapole trap held at a background pressure of about 10^{-3} Torr, where the trapping voltages can be adjusted to further enhance the formation of a specific CID ion. The ions were then transferred to the ICR cell, where the anion of interest (b_2) was isolated using a SWIFT excitation pulse[38]. Mass-isolated ions were irradiated with the tunable light from the free electron laser[34]. FELIX produces 5 μs long pulses with an energy of ~60 mJ (~30-40 mJ in the ICR cell) and a bandwidth of ~0.5% of the central wavelength. In the current experiments, the FELIX wavelength was scanned from 5.5 to 13 μm, which corresponds to a spectral range of 800-1800 cm^{-1}. Ions were irradiated for 3.8 s, while the entire FTICR sequence lasted 6 s. Four mass spectra were averaged at each wavelength point. Infrared multiple-photon dissociation (IR-MPD) occurs whenever the frequency of the laser is in resonance with a particular vibration in the ion[39-42]; tens to hundreds of photons are absorbed leading to the dissociation of the anion. Recording the photofragment yield as a function of wavelength gives an IR spectrum, where the yield is determined as $\sum I(\mathit{fragments}) / \sum I(\mathit{all\ ions})$.

Experimental spectra were compared to calculated spectra to identify the molecular structures. Optimized anion structures and computed spectra were generated by DFT using the B3LYP functional and the 6-31++G(d,p) basis set, implemented in Gaussian 03. This method was previously found to give a good prediction for the energies and IR spectra for deprotonated species [43,44]. Starting structures for the b_2 ions were constructed on the basis of those previously investigated by Harrison and coworkers[33] and using chemical intuition. Structures exhibiting resonance stabilization by conjugation and with a high number of H-bonding interactions were considered to be the most attractive candidates. The structures are presented in Figure 5.1. Computed harmonic frequencies are scaled by 0.97 and convoluted with a 25 cm^{-1} Gaussian line shape function to facilitate comparison with experimental spectra.

5.3. Results

The precursor peptides were selected on the basis of their activity and residue size. Tri-alanine was chosen because the alanine residue is nonpolar and relatively chemically inert to possible involvement of the side chain in the fragmentation reaction. Alanyl-tyrosyl-alanine possesses a bulky tyrosine residue which might influence the reaction by steric interactions. More importantly, the tyrosine hydroxyl group can deprotonate relatively easily (similar to phenol, pKa_{Tyr} ~10) and may thus compete for the negative charge with other acidic protons in the system.

5.3.1. Alanyl-Alanyl-Alanine

The b_2 fragment anion of deprotonated AAA at m/z 141 was generated by in-source CID, in particular by manipulating the settings of the cone voltage and the hexapole accumulation trap, and subsequently mass-isolated in the ICR cell. Upon IR activation the b_2 anion was found to undergo dissociation forming fragments with m/z 111, 113, 42 and 97 (ordered by their abundance), suggesting loss of formaldehyde (H_2CO), CO, methyl substituted oxazolone and CO_2, respectively. Observation of the CO_2 loss channel suggests that two oxygen atoms have become bonded to the same carbon atom in the b_2-fragment ion. This fragmentation pathway was also observed, though with CID of N-benzoylpeptides, in the work of Harrison and coworkers[30] and was interpreted as an indication for an oxazolone structure. These data provide circumstantial evidence for an oxazolone structure of the b_2 fragment.

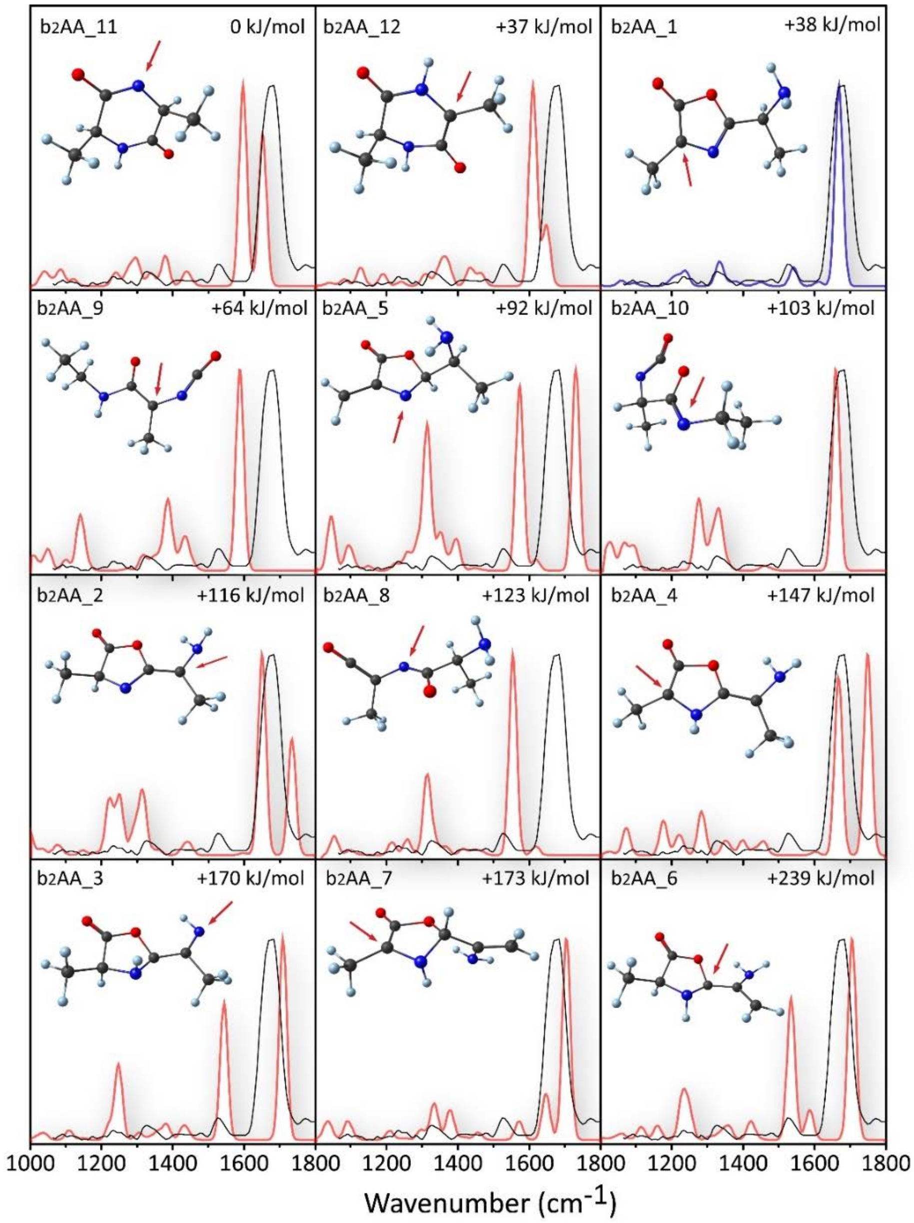


Figure 5.2. Experimental IRMPD spectrum of the b_2 fragment of deprotonated AAA (black line) compared with calculated spectra (red and blue dashed line) for various oxazolone, diketopiperazine and linear candidate structures (in order of increasing relative energy). The red arrow indicates the deprotonation site in each of the structures.

Figure 5.2 presents a direct comparison of the experimental spectrum with those computed for the various candidate structures of Figure 5.1. It is immediately obvious that only one of the suggested structures shows good agreement with the experimental spectrum, which is structure b_2AA_1, corresponding to a 2,4-substituted-5-oxazolone. The calculations formally place the negative charge for this structure on the α-carbon atom that has become integrated in the oxazolone ring. The dominant band observed in the IRMPD spectrum at 1670 cm^{-1} corresponds to the oxazolone C=O stretching mode. Since the structure is in fact an ester, one could have expected the carbonyl stretching mode to occur at slightly higher wavenumbers. However, the enol resonance induces a substantial red-shifting of this band. Bands with lower intensities appear at 1532, 1331, 1233 and 1096 cm^{-1} and correspond to N=C stretching, ring vibration, C-N in-ring stretching, and NH_2 twisting, respectively.

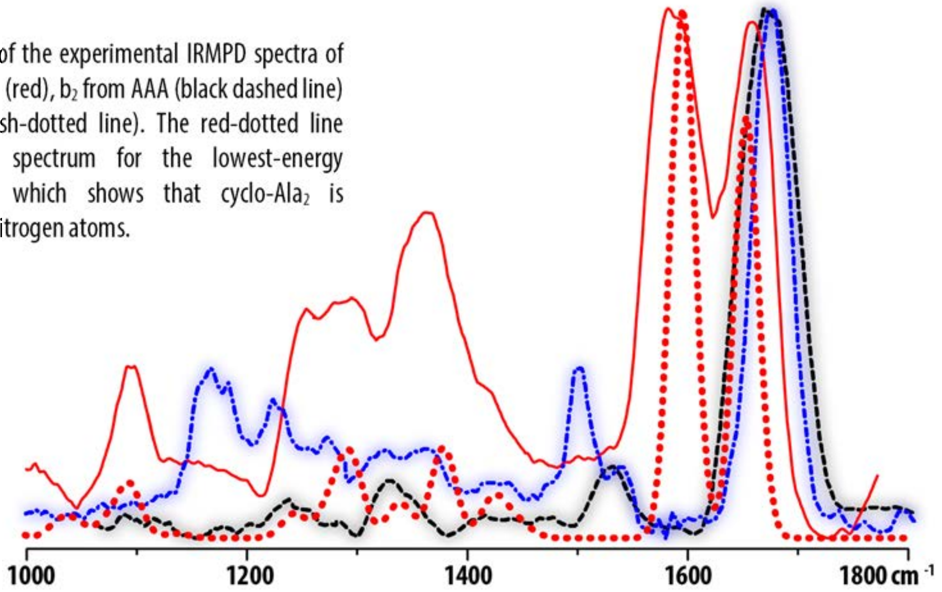
Calculated relative energies are given for each of the suggested structures in the upper left corner of each panel in Figure 5.2. Even though the spectrum of structure b_2AA_1 clearly fits best with the experimental spectrum, there are two structures with substantially lower energies. These are the diketopiperazine structures deprotonated on the ring nitrogen and carbon atoms, which are lower in energy than b_2AA_1 by 38 and 1 kJ/mol, respectively. It is obvious in the experimental spectrum that only one C=O stretching mode is present, while for the diketopiperazine structures two C=O stretching modes are predicted.

To further support our structural assignment and to exclude the presence of diketopiperazine structures, the spectrum of commercially available cyclic-dialanine, which has a diketopiperazine structure, was recorded. The results are shown in Figure 5.3.

The experimental spectrum of deprotonated cyclic-dialanine shows good agreement with the spectrum calculated for the N-deprotonated diketopiperazine structure (see Figure 5.3). The experimental spectrum indeed shows two carbonyl stretching modes, which occur at different frequencies because of the different molecular environment of the two carbonyl groups in the diketopiperazine ring. The band at 1660 cm^{-1} is due to the carbonyl stretch which is not affected by the amidate resonance, whereas the feature at 1590 cm^{-1} is due to the carbonyl involved in the amidate resonance, which gives partial single bond character to this CO bond.

It is further noticed that IRMPD dissociation channels are different from those observed for the b_2 fragment. Main fragment mass channels for cyclic-dialanine are m/z 42 and 97. While

Figure 5.3. Comparison of the experimental IRMPD spectra of deprotonated cyclo-dialanine (red), b_2 from AAA (black dashed line) and b_2 from AYA (blue dash-dotted line). The red-dotted line represents the calculated spectrum for the lowest-energy diketopiperazine structure, which shows that cyclo-Ala$_2$ is deprotonated on one of the nitrogen atoms.



the m/z 42 fragment may correspond to NCO^- or $NH=CH-CH_2^-$, an exact mass determination suggests the former stoichiometry. Mass 97 is due to the loss of a neutral mass 44 fragment, which could either be acetaldehyde (CH_3CHO) or carbon dioxide (CO_2); here the exact mass clearly suggests acetaldehyde loss, which is in line with experimental data by Bowie and coworkers on the dissociation of deprotonated 2,5-diketopiperazines indicating that CO_2-loss is unlikely[45]. In conclusion, both the spectrum as well as the fragmentation channels resulting from IRMPD are different for the AlaAla b_2 anion and the deprotonated cyclo-dialanine, providing independent experimental evidence against a diketopiperazine structure for the b_2 fragment.

5.3.2. Alanyl-Tyrosyl-Alanine

The b_2 fragment generated by front-end CID from deprotonated AYA is observed at m/z 233. Upon IR activation the b_2 anion was found to undergo dissociation into fragments at m/z 97, 93, 42 and 88 (ordered by their abundance). Mass 97 may correspond to the loss of neutral 4-methylen-2,5-cyclohexadiene-1-one followed by the loss of formaldehyde, while mass 93 likely corresponds to the formation of the phenoxide ($C_6H_5O^-$) anion. Both of these fragmentation channels are commonly observed for anionic peptides containing a tyrosine residue[46]. The identity of the m/z 88 fragment is not clear. Although its mass corresponds to that of deprotonated alanine, this would be unusual since the b_2 ion from AYA does not possess a carboxylate moiety, and moreover does not possess a C-terminal alanine residue. The analysis of the IRMPD spectrum of the b_2 fragment described below is independent of the identity of the fragment ions observed.

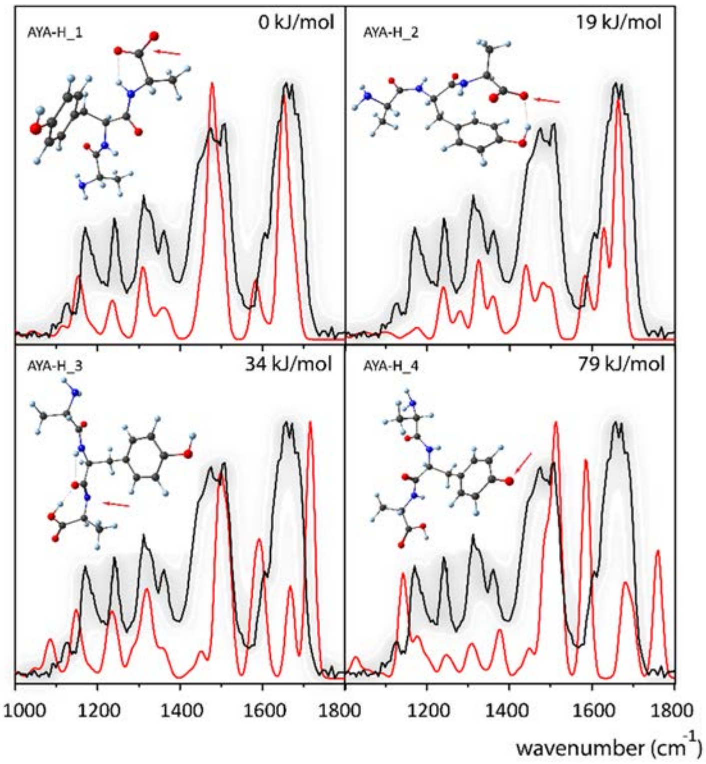


Figure 5.4. Comparison of theoretical and experimental spectra and energies (table) for $[AYA-H]^-$ for different deprotonation sites showing that the parent anion is carboxylate. Red arrow is pointed at the deprotonation site, though it should also be taken into account that there is actually a resonance for all of the presented structures i.e. deprotonation site is not absolutely localized.

Our main motivation for recording the IRMPD spectrum of the b_2 from AYA was to confirm the robustness of the deprotonated oxazolone motif found for the b_2 fragment from deprotonated AAA. Specifically, deprotonation may in this case occur on the tyrosine side chain forming a phenoxide anion. Hence, we can qualitatively compare the gas-phase acidity of the oxazolone motif with that of phenol. In addition, we recorded the IRMPD spectrum for the intact deprotonated AYA anion to establish the initial deprotonation site of the parent anion. Note that the deprotonation site of the Tyr amino acid has been the subject of substantial discussion recently[47-49]. Figure 5.4 unambiguously establishes a carboxylate structure for deprotonated AYA, analogous to what we found previously for deprotonated Tyr using IR spectroscopy[49].

The IRMPD spectrum recorded for the AYA b_2 fragment anion is shown in Figure 5.5, compared with calculated spectra for two oxazolone isomers, with deprotonation occurring either on the oxazolone α-carbon or on the phenol moiety of the Tyr side-chain. One can immediately see from the comparison that this b_2 fragment possesses a deprotonated oxazolone structure, analogous to the b_2 fragment of deprotonated AAA. Comparisons with spectra calculated for other possible structures can be found in the Ref[50]. The most intense absorption observed at 1675 cm^{-1} is attributed to the oxazolone C=O stretching mode, which appears slightly red-shifted at 1655 cm^{-1} in the calculated spectrum. Figure 5.3 shows that the experimentally observed frequency is consistent between the b_2 ions from AYA and AAA with a difference of only 5 cm^{-1}, suggesting that the discrepancy with theory may be an artifact of the harmonic frequency calculation. Assignment of the structure as a deprotonated oxazolone rather than a phenoxide anion is furthermore consistent with calculated relative free energies given for each of the structure in Figure 5.5.

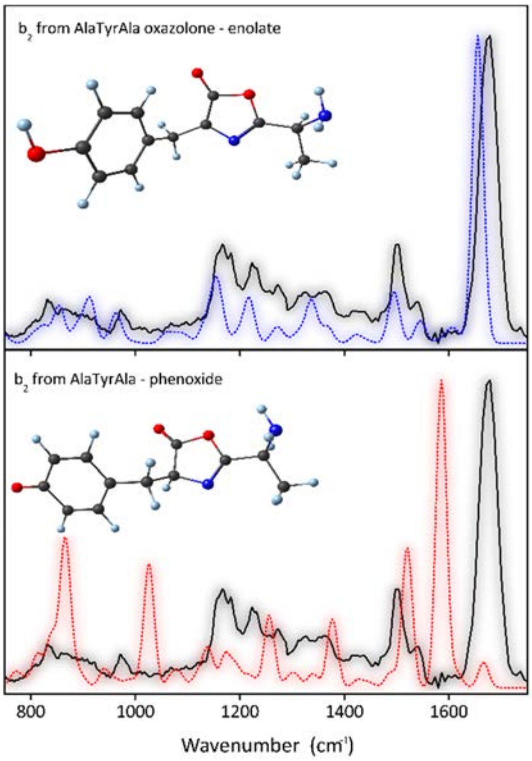


Figure 5.5. Experimental IRMPD spectrum of the b_2 fragment ion from deprotonated AYA (black in both panels) compared with calculated spectra for a deprotonated oxazolone structure (blue dots) and an oxazolone anion deprotonated on the Tyr residue (red dots).

Apart from the oxazolone carbonyl stretching mode, there are more similarities between the spectra of the b_2 ions from AYA and AAA. Absorption features observed at 1539, 1327 and 1231 cm^{-1} in the spectrum of b_2 from AYA can be correlated to absorptions at 1532, 1331, 1233 cm^{-1} in the spectrum of b_2 from AAA (see Figure 5.3). According to the calculations, these peaks are due to the C=N stretching vibration, the oxazolone ring vibration and the oxazolone C-N stretching mode, respectively. In addition, features observed at 833, 809, 1080 and 1168 cm^{-1} in the experimental spectrum of b_2 from AYA correspond to bands at 854, 824, 1080 and 1161 cm^{-1}, respectively, in the theoretical spectrum and are due to aromatic out-of-plane and in-plane C-H vibrations. Furthermore, a band of special interest located at 1153 cm^{-1} in the calculations is observed as a shoulder at 1156 cm^{-1} in the experiment. This peak is assigned as the hydroxyl COH bending mode of the phenol moiety and is very close in frequency to the analogous mode in the IRMPD spectrum of deprotonated tyrosine[49]. This similarity provides further experimental evidence for the tyrosine side-chain not being deprotonated. In addition, the feature at 1223 cm^{-1} (1218 cm^{-1} in the theoretical spectrum) is due to the phenolic C-O stretching mode. Finally, we note that the NH_2 bending vibration calculated at 912 cm^{-1} is not well reproduced in the experimental spectrum. Previous IRMPD spectra of deprotonated amino acids have shown that the frequency and intensity of this mode are often found to deviate strongly from their theoretical values [49], which we suspect to be due to the very anharmonic potential for this vibration[51,52]. All experimental and theoretical data thus strongly suggest that both deprotonated peptides form b_2 fragments with an oxazolone structure and that deprotonation occurs on the α-carbon embedded in the oxazolone ring.

5.4. Discussion

One may now wonder what makes this particular oxazolone structure more attractive than all other candidate structures presented in Figure 5.1. As described above, the oxazolone ring is believed to be formed by nucleophilic attack of the N-terminal amide oxygen onto the neighboring amide carbon (see Scheme 5.2)[29], whereas the thermodynamically more favored diketopiperazine structure would have to be formed by nucleophilic attack of the N-terminal nitrogen onto the second amide carbon. This latter reaction requires a *trans* to *cis* isomerisation of the N-terminal amide bond before cleavage of the peptide bond occurs. For protonated peptides, rotation around the amide bond faces a relatively low energetic penalty due to the reactive configuration in which the mobile proton has migrated to the amide nitrogen, reducing the bond order of the amide bond. However, while energetically feasible, this isomerization is entropically unfavorable[53-57], preventing the formation of diketopiperazine structures. Although no transition state calculations were performed for deprotonated peptides, in addition to possible entropic constraints, *trans*-to-*cis* isomerisation likely faces higher energetic barriers, since in the reactive configuration where the amide moiety is deprotonated (Scheme 5.2), the double-bond character of the amide

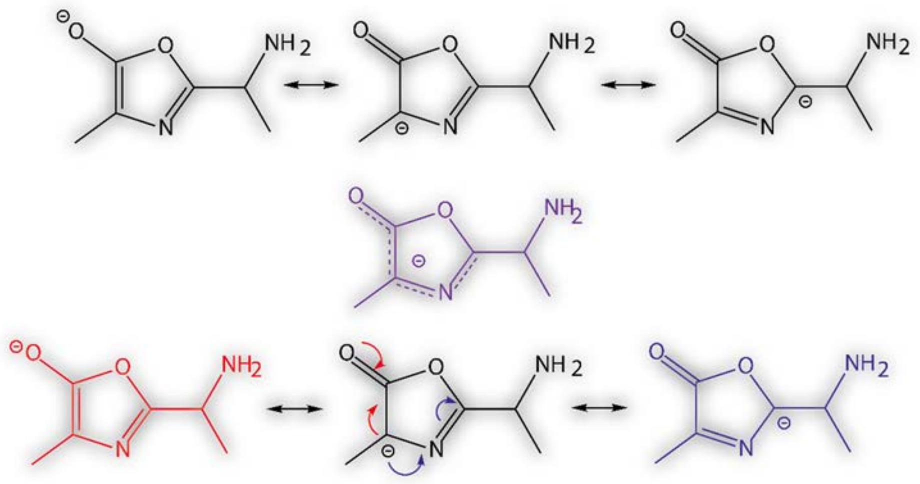


Figure 5.6. Resonant structures for the b$_2$AA_1 oxazolone structure. The negative charge is distributed over five heavy atoms providing particular stability to this oxazolone structure. The overall conjugated structure is presented in violet.

bond is enhanced rather than reduced. For oxazolone structures to be formed on the other hand, only rotation around the N-C$_\alpha$ bond is needed. Furthermore, formation of the diketopiperazine structure may require deprotonation of the N-terminal amino group for it to become sufficiently nucleophilic, which is energetically unfavorable in the gas phase as it would strongly localize the negative charge. Trans-to-cis isomerization would then have to occur around a neutral, conjugated amide bond facing a substantial thermodynamic barrier. These considerations make formation of the diketopiperazine an unlikely reaction pathway and provide qualitative arguments for fragmentation pathways leading to oxazolone structures.

Of all possible oxazolone structures in Figure 5.1, the structure identified in the experiments has the lowest energy and here we rationalize why this structure is particularly favorable. A close inspection of this structure in Figure 5.6 reveals that there is a high level of charge delocalization in this oxazolone b$_2$ structure, in which the negative charge can be delocalized over five heavy atoms. None of the other oxazolone forms in Figure 5.1 allows one to draw more than just two reasonable resonance structures. This conjecture can be further supported with calculated bond lengths of the b$_2$AA_1 deprotonated structure as compared with those for its neutral counterpart (Figure 5.7). Elongation of the C$_a$=O bond by 0.046 Å and shortening of the C$_a$-C$_c$ bond by 0.126 Å is a consequence of the enolate resonance (red isomer in Figure 5.6). On the other hand, substantial contraction of the C$_c$-N bond occurs (0.075 Å) as well as an elongation of the N=C$_b$ bond (0.029 Å) as a result of the imine resonance (shown as the blue isomer in Figure 5.6). This results in a conjugated structure for the b$_2$AA_1 anion as presented in violet in Figure 5.6.

Finally, we note that the fragment ion structures that have been identified in this and a previous study[21] on the fragmentation of deprotonated tripeptides support the conclusions of Harrison, Siu and coworkers[29,32,33]. In particular, we showed that the a$_3$ fragment of deprotonated AAA has an amidate structure that corresponds to the suggested structure and the fragmentation pathway of Harrison shown here in Scheme 5.3. Our results for the

b_2 fragment anions described in the current study are in support of Harrison's suggested *b*-ion structure, possibly formed via the a_3 anion (see Scheme 5.3).

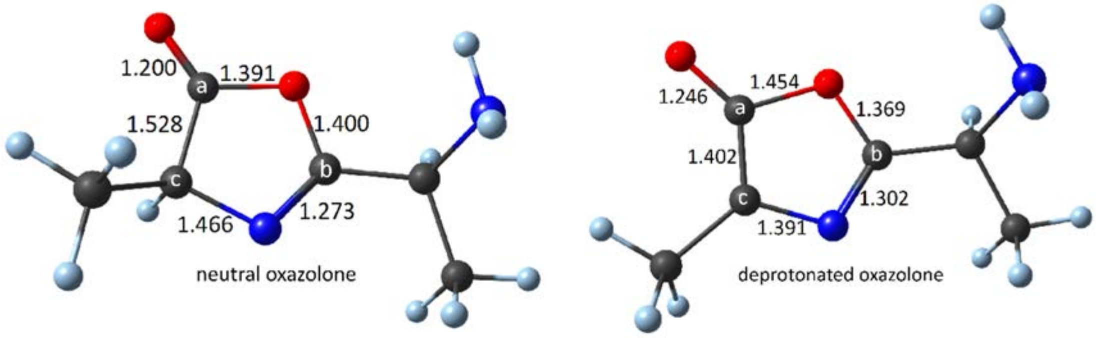


Figure 5.7. Bond lengths in deprotonated and neutral oxazolone ring

5.5. Conclusions

The first spectroscopic investigation of anionic *b*-ions has been presented; gas-phase IRMPD spectra for the b_2 fragments of the deprotonated peptides AAA and AYA have been recorded. Comparisons with calculated spectra show unambiguously that the b_2 fragments generated from both peptides have an oxazolone structure, which is deprotonated on the α-carbon in the oxazolone ring. No evidence was found for b_2 fragments having the lower-energy diketopiperazine structure, which was further confirmed by recording an experimental spectrum for deprotonated cyclo-dialanine that is substantially different from that recorded for the CID generated b_2 ion of deprotonated AAA. The deprotonated oxazolone *b*-ion structure identified here is shown to be a robust motif in the sense that it is also formed in the presence of other acidic protons, in this case on the tyrosine side-chain.

Formation of the deprotonated oxazolone structure is likely due to an energetically and/or entropically favorable transition state as compared to that for the formation of the diketopiperazine structure. The latter pathway requires a *trans*-to-*cis* isomerisation of the peptide bond, which possibly represents a higher energetic barrier as compared to an analogous isomerisation in protonated peptides, although computational investigations are required to substantiate this qualitative argument. Furthermore, of all possible oxazolone structures that can be formed, the one where deprotonation occurs on the oxazolone C_α is particularly stable due to a high degree of charge delocalization, which is an important factor in the gas-phase stability of ions.

In more general terms, this study shows that characterization of CID peptide fragment structures by IRMPD spectroscopy is also successful for anionic species. A deeper understanding of peptide fragmentation in negative ion mode may eventually lead to an increased application in peptide sequencing strategies.

Acknowledgment

We would like to thank Britta Redlich, Lex van der Meer, Giel Berden and others at the FELIX facility for their skillful assistance. This work is part of the research program of FOM, which is financially supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). JO thanks the Stichting Physica for support.

REFERENCES

(1) McLafferty, F. W. *Tandem mass spectrometry (MS/MS): a promising new analytical technique for specific component determination in complex mixtures* Acc. Chem. Res. **1980**, *13*, 33.
(2) Siuzdak, G. *The emergence of mass-spectrometry in biochemical research* Proc. Natl. Acad. Sci. U. S. A. **1994**, *91*, 11290.
(3) Yates, J. R. *Mass spectrometry and the age of the proteome* J. Mass Spectrom. **1998**, *33*, 1.
(4) Biemann, K. *Contributions of mass spectrometry to peptide and protein structure* Biol. Mass Spectrom. **1988**, *16*, 99.
(5) Harrison, A. G. *Peptide Sequence Scrambling Through Cyclization of b5 Ions* J. Am. Soc. Mass Spectrom. **2008**, *19*, 1776.
(6) Bleiholder, C.; Osburn, S.; Williams, T. D.; Suhai, S.; Van Stipdonk, M.; Harrison, A. G.; Paizs, B. *Sequence-Scrambling Fragmentation Pathways of Protonated Peptides* J. Am. Chem. Soc. **2008**, *130*, 17774.
(7) Yu, L.; Tan, Y.; Tsai, Y.; Goodlett, D. R.; Polfer, N. C. *On the relevance of peptide sequence permutations in shotgun proteomics studies* J. Proteome Res. **2011**, *10*, 2409.
(8) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *Spectroscopic and Theoretical Evidence for Oxazolone Ring Formation in Collision-Induced Dissociation of Peptides* J. Am. Chem. Soc. **2005**, *127*, 17154.
(9) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *Infrared spectroscopy and theoretical studies on gas-phase protonated leu-enkephalin and its fragments: Direct experimental evidence for the mobile proton* J. Am. Chem. Soc. **2007**, *129*, 5887.
(10) Savitski, M. M.; Falth, M.; Fung, Y. M. E.; Adams, C. M.; Zubarev, R. A. *Bifurcating Fragmentation Behavior of Gas-Phase Tryptic Peptide Dications in Collisional Activation* J. Am. Soc. Mass Spectrom. **2008**, *19*, 1755.
(11) Oomens, J.; Young, S.; Molesworth, S.; van Stipdonk, M. *Spectroscopic Evidence for an Oxazolone Structure of the* $b_{(2)}$ *Fragment Ion from Protonated Tri-Alanine* Journal of the American Society for Mass Spectrometry **2009**, *20*, 334.
(12) Yoon, S. H.; Chamot-Rooke, J.; Perkins, B. R.; Hilderbrand, A. E.; Poutsma, J. C.; Wysocki, V. H. *IRMPD Spectroscopy Shows That AGG Forms an Oxazolone* b_2^+ *Ion* J. Am. Chem. Soc. **2008**, *130*, 17644.
(13) Perkins, B. R.; Chamot-Rooke, J.; Yoon, S. H.; Gucinski, A. C.; Somogyi, A. r. d.; Wysocki, V. H. *Evidence of Diketopiperazine and Oxazolone Structures for HA* b_2^+ *Ion* J. Am. Chem. Soc. **2009**, *131*, 17528.
(14) Molesworth, S.; Osburn, S.; Van Stipdonk, M. *Influence of Size on Apparent Scrambling of Sequence During CID of b-Type Ions* J. Am. Soc. Mass Spectrom. **2009**, *20*, 2174.
(15) Molesworth, S.; Osburn, S.; Van Stipdonk, M. *Influence of Amino Acid Side Chains on Apparent Selective Opening of Cyclic* $b_{(5)}$ *Ions* J. Am. Soc. Mass Spectrom. **2010**, *21*, 1028.
(16) Chen, X.; Yu, L.; Steill, J. D.; Oomens, J.; Polfer, N. C. *Effect of Peptide Fragment Size on the Propensity of Cyclization in Collision-Induced Dissociation: Oligoglycine b(2)-b(8)* J. Am. Chem. Soc. **2009**, *131*, 18272.
(17) Chen, X. A.; Steill, J. D.; Oomens, J.; Polfer, N. C. *Oxazolone Versus Macrocycle Structures for Leu-Enkephalin* $b_{(2)}$*-*$b_{(4)}$*: Insights from Infrared Multiple-Photon Dissociation Spectroscopy and Gas-Phase Hydrogen/Deuterium Exchange* J. Am. Soc. Mass Spectrom., *21*, 1313.
(18) Erlekam, U.; Bythell, B. J.; Scuderi, D.; Van Stipdonk, M.; Paizs, B.; Maitre, P. J. Am. Chem. Soc. **2009**, *131*, 11503.
(19) Verkerk, U. H.; Siu, C.-K.; Steill, J. D.; El Aribi, H.; Zhao, J.; Rodriquez, C. F.; Oomens, J.; Hopkinson, A. C.; Siu, K. W. M. a_2 *Ion Derived from Triglycine: An N1-Protonated 4-Imidazolidinone* J. Phys. Chem. Lett. **2010**, *1*, 868.

(20) Bythell, B. J.; Maitre, P.; Paizs, B. *Cyclization and Rearrangement Reactions of a_n Fragment Ions of Protonated Peptides* J. Am. Chem. Soc. **2010**, *132*, 14766.

(21) Oomens, J.; Steill, J. D. *The Structure of Deprotonated Tri-Alanine and Its a3- Fragment Anion by IR Spectroscopy* J. Am. Soc. Mass Spectrom. **2010**, *21*, 698.

(22) Chass, G. A.; Marai, C. N. J.; Harrison, A. G.; Csizmadia, I. G. *Fragmentation reactions of $a_{(2)}$ ions derived from deprotonated dipeptides - A synergy between experiment and theory* J. Phys. Chem. A **2002**, *106*, 9695.

(23) Kulik, W.; Heerma, W. *Fast Atom Bombardment Tandem Mass-Spectrometry for Amino-Acid Sequence Determination in Tripeptides* Biomed. Environ. Mass Spectrom. **1989**, *18*, 910.

(24) Van Setten, D.; Kulik, W.; Heerma, W. *Isomeric tripeptides: A study on structure-spectrum relationship* Biol. Mass Spectrom. **1990**, *19*, 475.

(25) Bache, N.; Rand, K. D.; Roepstorff, P.; Ploug, M.; Jørgensen, T. J. D. *Hydrogen Atom Scrambling in Selectively Labeled Anionic Peptides Upon Collisional Activation by MALDI Tandem Time-of-Flight Mass Spectrometry* J. Am. Soc. Mass Spectrom. **2008**, *19*, 1719.

(26) Eckersley, M.; Bowie, J. H.; Hayes, R. N. *Collision-Induced Dissociations of Deprotonated Peptides, Dipeptides and Tripeptides with Hydrogen and Alkyl a-Groups—an Aid to Structure Determination* Org. Mass Spectrom. **1989**, *24*, 597.

(27) Waugh, R. J.; Bowie, J. H. *A review of the collision induced dissociations of deprotonated dipeptides and tripeptides. An aid to structure determination.* Rapid Commun. Mass Spectrom. **1994**, *8*, 169.

(28) Bowie, J. H.; Brinkworth, C. S.; Dua, S. *Collision-induced fragmentations of the $(M-H)^{(-)}$ parent anions of underivatized peptides: An aid to structure determination and some unusual negative ion cleavages* Mass Spectrometry Reviews **2002**, *21*, 87.

(29) Harrison, A. G.; Siu, K. W. M.; El Aribi, H. *Amide bond cleavage in deprotonated tripeptides: a newly discovered pathway to " $b_{(2)}$ ions* Rapid Commun. Mass Spectrom. **2003**, *17*, 869.

(30) Harrison, A.; Young, A. *Fragmentation of deprotonated N-benzoylpeptides: Formation of deprotonated oxazolones* J. Am. Soc. Mass Spectrom. **2004**, *15*, 446.

(31) Harrison, A. G. *Sequence-specific fragmentation of deprotonated peptides containing H or alkyl side chains* J. Am. Soc. Mass Spectrom. **2001**, *12*, 1.

(32) Harrison, A. G.; Young, A. B. *Fragmentation reactions of deprotonated peptides containing proline. The proline effect* J. Mass Spectrom. **2005**, *40*, 1173.

(33) Chass, G. A.; Marai, C. N. J.; Setiadi, D. H.; Csizmadia, I. G.; Harrison, A. G. *A Hartree-Fock, MP2 and DFT computational study of the structures and energies of " b_2 ions derived from deprotonated peptides. A comparison of method and basis set used on relative product stabilities* J. Mol. Struct. Theochem. **2004**, *675*, 149.

(34) Oepts, D.; van der Meer, A. F. G.; van Amersfoort, P. W. *The Free-Electron-Laser user facility FELIX* Infrared Phys. Technol. **1995**, *36*, 297.

(35) Polfer, N. C.; Oomens, J. *Reaction products in mass spectrometry elucidated with infrared spectroscopy* Physical Chemistry Chemical Physics **2007**, *9*, 3804.

(36) Valle, J. J.; Eyler, J. R.; Oomens, J.; Moore, D. T.; van der Meer, A. F. G.; von Helden, G.; Meijer, G.; Hendrickson, C. L.; Marshall, A. G.; Blakney, G. T. *Free electron laser-Fourier transform ion cyclotron resonance mass spectrometry facility for obtaining infrared multiphoton dissociation spectra of gaseous ions* Rev. Sci. Instrum. **2005**, *76*, 023103.

(37) Mize, T. H.; Taban, I.; Duursma, M.; Seynen, M.; Konijnenburg, M.; Vijftigschild, A.; Doornik, C. V.; Rooij, G. V.; Heeren, R. M. A. *A modular data and control system to improve sensitivity, selectivity, speed of analysis, ease of use, and transient duration in an external source FTICR-MS* Int. J. Mass Spectrom. **2004**, *235*, 243.

(38) Marshall, A. G.; Wang, T. C. L.; Ricca, T. L. *Tailored excitation for Fourier transform ion cyclotron mass spectrometry* J. Am. Chem. Soc. **1985**, *107*, 7893.

(39) Polfer, N. C.; Oomens, J. *Vibrational spectroscopy of bare and solvated ionic complexes of biological relevance* Mass Spectrom. Rev. **2009**, *28*, 468.

(40) Oomens, J.; Sartakov, B. G.; Meijer, G.; von Helden, G. *Gas-phase infrared multiple photon dissociation spectroscopy of mass-selected molecular ions* Int. J. Mass Spectrom. **2006**, *254*, 1

(41) Polfer, N. C. *Infrared multiple photon dissociation spectroscopy of trapped ions* Chem. Soc. Rev. **2011**, *40*, 2211.

(42) MacAleese, L.; Maitre, P. *Infrared Spectroscopy of Organometallic Ions in the Gas Phase: From Model to Real World Complexes* Mass Spectrom. Rev. **2007**, *26*, 583.

(43) Merrill, G. N.; Kass, S. R. *Calculated Gas-Phase Acidities Using Density Functional Theory: Is It Reliable?* J. Phys. Chem. **1996**, *100*, 17465.

(44) Steill, J. D.; Oomens, J. *Action Spectroscopy of Gas-Phase Carboxylate Anions by Multiple Photon IR Electron Detachment/Attachment* J. Phys. Chem. A **2009**, *113*, 4941.

(45) Wabnitz, P. A.; Waugh, R. J.; Eckersley, M. A.; Dua, S.; Blumenthal, T.; Bowie, J. H. *The negative ion mass spectra of deprotonated 2,5-diketopiperazines* International Journal of Mass Spectrometry and Ion Processes **1996**, *154*, 193.

(46) Waugh, R. J.; Bowie, J. H.; Hayes, R. N. *Collision induced dissociations of deprotonated peptides: Dipeptides containing phenylalanine, tyrosine, histidine and tryptophan* Int. J. Mass Spectrom. Ion Process. **1991**, *107*, 333.

(47) Jones, C. M.; Bernier, M.; Carson, E.; Colyer, K. E.; Metz, R.; Pawlow, A.; Wischow, E. D.; Webb, I.; Andriole, E. J.; Poutsma, J. C. *Gas-phase acidities of the 20 protein amino acids* Int. J. Mass Spectrom. **2007**, *267*, 54.

(48) Tian, Z.; Wang, X.-B.; Wang, L.-S.; Kass, S. R. *Are Carboxyl Groups the Most Acidic Sites in Amino Acids? Gas-Phase Acidities, Photoelectron Spectra, and Computations on Tyrosine, p-Hydroxybenzoic Acid, and Their Conjugate Bases* J. Am. Chem. Soc. **2009**, *131*, 1174.

(49) Oomens, J.; Steill, J. D.; Redlich, B. *Gas-Phase IR Spectroscopy of Deprotonated Amino Acids* J. Am. Chem. Soc. **2009**, *131*, 4310.

(50) Grzetic, J.; Oomens, J. *Spectroscopic evidence for an oxazolone structure in anionic b-type peptide fragments* J. Am. Soc. Mass Spectrom. **2012**, *23*, 290.

(51) Sinclair, W. E.; Pratt, D. W. *Structure and vibrational dynamics of aniline and aniline–Ar from high resolution electronic spectroscopy in the gas phase* J. Chem. Phys. **1996**, *105*, 15.

(52) Piest, H.; Helden, v. G.; Meijer, G. *Infrared spectroscopy of jet-cooled neutral and ionized aniline-Ar* J. Chem. Phys. **1999.**, *110*, 2010.

(53) Paizs, B.; Suhai, S. *Towards understanding the tandem mass spectra of protonated oligopeptides. 1: Mechanism of amide bond cleavage* J. Am. Soc. Mass Spectrom. **2004**, *15*, 103.

(54) Paizs, B.; Suhai, S. *Combined quantum chemical and RRKM modeling of the main fragmentation pathways of protonated GGG. II. Formation of* $b_{(2)}$, $y_{(1)}$, *and* $y_{(2)}$ *ions* Rapid Commun. Mass Spectrom. **2002**, *16*, 375.

(55) Paizs, B.; Lendvay, G.; Vekey, K.; Suhai, S. *Formation of* $b_{(2)}^{(+)}$ *ions from protonated peptides: an ab initio study* Rapid Commun. Mass Spectrom. **1999**, *13*, 525.

(56) Balta, B.; Aviyente, V.; Lifshitz, C. *Elimination of water from the carboxyl group of* $GlyGlyH^+$ J. Am. Soc. Mass Spectrom. **2003**, *14*, 1192.

(57) Armentrout, P. B.; Heaton, A. L. *Thermodynamics and mechanisms of protonated diglycine decomposition: a computational study* J Am Soc Mass Spectrom **2012**, *23*, 621.

Chapter 6.

Effect of the Asn side chain on the dissociation of deprotonated peptides elucidated by IRMPD spectroscopy

Infrared ion spectroscopy using the free electron laser FELIX was applied to identify the structure of *b*-type peptide fragments generated by collision and IR multiple-photon induced dissociation from singly deprotonated peptides containing an asparagine residue, in particular AlaAsnAla (ANA) and AlaAlaAsnAla (AANA). IR spectra were recorded over the 800-1800 cm^{-1} spectral range by multiple-photon dissociation (IRMPD) spectroscopy and have been compared with density functional theory (DFT) calculated spectra at the B3LYP/6-31++G(d,p) level for different isomeric ion structures for structural characterization. Results unambiguously show that the b_2 and b_3 fragment anions do not possess the common oxazolone or diketopiperazine structure, but involve cyclization of the asparagine side chain. Nucleophilic attack from the side chain amide nitrogen on the peptide backbone carbonyl carbon leads to the formation of cyclic succinimide structures. Deprotonation is shown to occur on the succinimide nitrogen, which delocalizes the negative charge over two adjacent carbonyl groups thus enhancing the gas-phase stability. *

---

* Adapted from: J.Grzetic and J.Oomens, *Effect of the Asn side chain on the dissociation of deprotonated peptides elucidated by IRMPD spectroscopy*, Int. J. of Mass Spectrom. (2013), 354–355, 70-77

6.1. Introduction

Sequencing of proteolytic peptides using mass spectrometric techniques is one of the main tools in proteomics research[1-3]. It is essentially used to identify amino acid sequences of peptides. The practical use is based on the empirical differences in the m/z values of peptide fragments created by collision-, photon- or electron-induced dissociation of protonated peptides. While this method of analyzing protein structures is widely accepted, the fundamental chemical reactions underlying peptide dissociation processes are only partially understood[4]. Using chemical intuition and the results of tandem MS experiments alone, various peptide fragment structures and corresponding dissociation pathways have been suggested in the literature. However, further computational and experimental evidence is usually required to firmly establish the suggested reaction pathways and product structures, especially since various isomeric and tautomeric structures are often conceivable. In addition to information based purely on quantum-chemical calculations, experimental evidence for fragment structures has recently been provided via a combination of MS instrumentation and tunable infrared lasers, which allows one to record IR spectra of the peptide fragment ions of interest[5-7]. Combining these experimental IR spectra with results of computational chemistry (or with experimental spectra of model compounds) the structure of peptide fragment ions can often be identified with high confidence[8,9].

Spectroscopic studies of collision induced dissociation (CID) products of protonated peptides have mainly focused on *a*- and *b*-type fragments (peptide fragment nomenclature adopted from Roepstorff[10]). It has been shown that depending on the amino acid sequence, b_2-type fragments can possess oxazolone or diketopiperazine structures[11-13]. Very recently we have also identified residue specific fragmentation of protonated peptides containing glutamine or asparagine residues[14]. In contrast to formation of oxazolone and diketopiperazine b_2-type structures, here we noticed the interference of Asn and Gln residues in *b*-type ion formation; dissociation occurs via nucleophilic attack of the side chain amide nitrogen atom of Asn and Gln on the adjacent peptide carbonyl carbon, which leads to formation of succinimide and glutarimide ring structures for these *b*-type ions, respectively. In addition to the fragment structures above, ion spectroscopy was also instrumental in the discovery that larger *b*-ions can undergo “head-to-tail” cyclization leading to macrocyclic structures, which may reopen generating fragments with a permuted sequence[15-19]. Formation of macrocyclic structures was also reported for *a*-type peptide fragments[20,21].

While MS-based peptide sequencing in proteomics research is almost exclusively based on fragmentation of protonated peptides, analyzing the fragmentation of deprotonated peptides may generate different fragmentation patterns and thus increase sequence coverage[22-25]. Moreover, for peptides containing many acidic residues or for phosphorylated peptides, running in negative ion mode, i.e., sequencing the deprotonated peptide, may be more efficient[26]. While fragmentation of protonated peptides mainly generates *a*-, *b*- and *y*-type backbone fragments, peptide anion fragmentation features in addition *c*- and *z*-type ions, which may also provide additional sequence information. Various studies have therefore addressed the mechanisms of fragmentation of deprotonated peptides using tandem MS and computational methods[23,25,27-35]. More recently, IR spectroscopy has been applied to elucidate the structure of anionic *a*-, *b*- and *c*-type fragments; for the a_3 anion generated from deprotonated Ala_3 a linear amidate structure was established[36]. The b_2-fragments generated from deprotonated AlaTyrAla and Ala_3 adopt oxazolone structures deprotonated at the oxazolone carbon atom adjacent to the carbonyl carbon[37]. c-Type anions were spectroscopically shown to possess linear structures, where the deprotonation site is either the peptide bond nitrogen atom or the side chain for residues with relatively high gas-phase acidity[38].

As established for protonated peptides, the side chain of Asn and Gln may become involved in the rearrangement reactions forming b_2-type fragments[39-42]. It is therefore of interest to investigate what the effect of these residues is in the dissociation of deprotonated peptides. Glutamine and asparagine are amino acids which feature an amide group, which acts as a relatively strong nucleophile. Furthermore, they have been addressed as uniquely unstable residues under physiological conditions (*i.e.* in aqueous solution), spontaneously undergoing protein deamidation, where the amide group of Asn or Gln is converted to a carboxylate moiety. Structural changes of the protein upon such reactions induce biological responses such as molecular clocks for the timing of biological processes or acceleration of amyloid aggregate formation (pathological protein formation)[43,44].

Here we present an IRMPD spectroscopic study on *b*-type peptide fragments generated from deprotonated peptides containing Asn, in particular for the b_2 anion from AlaAsnAla and the b_3 anion from AlaAlaAsnAla. A structural assignment is proposed based upon comparison of the experimental spectra with theoretically generated IR spectra for various isomeric candidate fragment structures.

6.2. Experimental and computational methods

6.2.1. IRMPD Spectroscopy

Infrared spectra were obtained by infrared multiple-photon dissociation (IRMPD) spectroscopy using the Fourier transform ion cyclotron resonance (FTICR) mass spectrometer coupled to the beam-line of the infrared free-electron laser FELIX[7,45,46].

Deprotonated peptides were generated by electrospray ionization (ESI, Micromass Z-spray source) from approximately 0.5 mmol solutions of the sample peptide in a 1:1

acetonitrile/water mixture with a drop of 10 mmol ammonium hydroxide added to enhance the deprotonation. Solutions were always freshly made in order to avoid possible chemical modifications of the samples. Peptides (AlaAsnAla and AlaAlaAsnAla) were purchased from Genecust (Luxemburg) and used without further purification. The b_2 fragment from deprotonated ANA was generated by collision induced dissociation (CID) in the high-pressure region at the inlet of the mass spectrometer. The b_3 anion from AANA was generated in the FTICR cell by 0.1 s long CO_2 laser (Universal laser systems, Scottsdale, USA, ULR-25) irradiation. The choice between these two methods of fragmentation was based on their efficiency in generating a high ion signal on the fragment ion of interest. For systems where both methods were used, we found no evidence of their IR spectra being different, suggesting that the same structure is generated independent of the fragmentation method used.

Inside the ICR cell, the ion of interest was isolated using a stored waveform inverse Fourier transform (SWIFT) excitation pulse[47]. While stored inside the ICR cell, the mass-selected fragment ion was irradiated with the IR beam from FELIX. FELIX produces 5 μs long pulses with an output pulse energy (at the ICR cell) of about 30-40mJ, while the bandwidth is about 0.5% of the central wavelength. The scanning range in these experiments was approximately 800-1900 cm^{-1}. Ions were irradiated for 2.5 s while the entire FTICR sequence was 4.5 s long.

IRMPD spectroscopy is based on the assumption that only photons resonant with a vibrational band of the molecular ion are absorbed. The energy of the absorbed photon is randomized over all internal degrees of freedom by intramolecular vibrational distribution (IVR) on a sub-ps timescale. Thus, tens to hundreds of photons can be absorbed during a single FEL pulse raising the internal energy of the ion until it undergoes unimolecular dissociation. (See Ref.[48] for an interesting recent discussion by Detlef Schröder and his coworkers on the effects of IVR in IRMPD spectroscopy and possible breakdown of the assumption of ergodicity). After irradiation, a mass spectrum is generated using an excite/detect sequence as implemented in the FTICR software[49]. Three mass spectra are collected at each wavelength point and averaged. An IR spectrum is then generated by plotting the fragment ion yield as a function of wavelength, where the yield is calculated as $\sum I(\text{fragments})/\sum I(\text{all ions})$. The yield is linearly corrected for the laser power profile over the scan range.

6.2.2. Density functional theory calculations

DFT calculations were performed using the B3LYP functional with the 6-31++G(d,p) basis set implemented in Gaussian09[50]. This method was found to give satisfactory results in previous experiments considering deprotonated species[51,52]. Candidate input structures for these calculations are based on chemical intuition, where the number of hydrogen bonds was maximized and cis configurations of peptide bonds as well as eclipsed steric interactions were avoided. Computed harmonic vibrational frequencies were scaled by 0.975 and convoluted with a 25 cm^{-1} fwhm Gaussian line shape function to facilitate comparison with experimental spectra.

6.3. Results and discussion

6.3.1. b_2 anion from AlaAsnAla

As for protonated peptides, amide bond cleavage is a common peptide fragmentation pattern for deprotonated peptides, leading to the formation of *b*- and *y*-type ions. Formation of *b*-type fragments from deprotonated peptides was first reported by Heerma and co-workers[53]. Two different b_2-fragment ion structures were proposed: a linear amidate *b*-ion with a C=C=O unit at the C-terminus[54-56] and a deprotonated oxazolone structure[34,57]. The b_2 fragments generated from deprotonated Ala_3 and deprotonated AlaTyrAla have been spectroscopically identified as oxazolones, as had indeed been suggested by Harrison et al.[34].

Apart from an oxazolone structure, conceivable structures for the b_2-ion of deprotonated AlaAsnAla are diketopiperazine, imino-γ-butyrolactone and succinimide as shown in Figure 7.1. While the oxazolone and diketopiperazine structures can be formed through nucleophilic attack by backbone O- and N-atoms, the imino-γ-butyrolactone and succinimide structures are formed through nucleophilic attack by the Asn side chain O- and N-atoms, respectively. For protonated peptides containing Asn, these alternative b-ion structures have been suggested[39] and the succinimide has indeed been spectroscopically identified very recently[14].

Upon CID, deprotonated AlaAsnAla was found to dissociate mainly into the b_2 ion at m/z 184, with a minor y_1 fragment at m/z 88. Upon IRMPD, the isolated b_2 anion dissociates into m/z 96, m/z 113 and m/z 87 (in order of their abundance), corresponding to C_4H_2NO (maleimide anion, as will be shown below), deprotonated 5-amino-succinimide and the c_1 fragment, respectively. This fragmentation pattern is unlike that observed for oxazolone b_2-ions (Figure 6.1, structure III) of deprotonated Ala_3 and AlaTyrAla, for which the main neutral loss channels are H_2CO, CO_2 and CO[37,57]. Moreover, the fragmentation pattern observed here is clearly different from the dissociation pathways of deprotonated diketopiperazines (Figure 6.1, structure II)[58]. Analyzing only this MS/MS part of the experimental data suggests that the fragments generated by IRMPD of the ANA b_2 anion are consistent with structures where the first peptide bond as well as the second N-C_α bond do not belong to any cyclic moieties, such as for structures I, IV and V in Figure 6.1. However, based on this data alone both succinimide and butyrolactone structures remain possible, since both of these structures match the hypothetical MS "puzzle" as depicted in Figure 6.2.

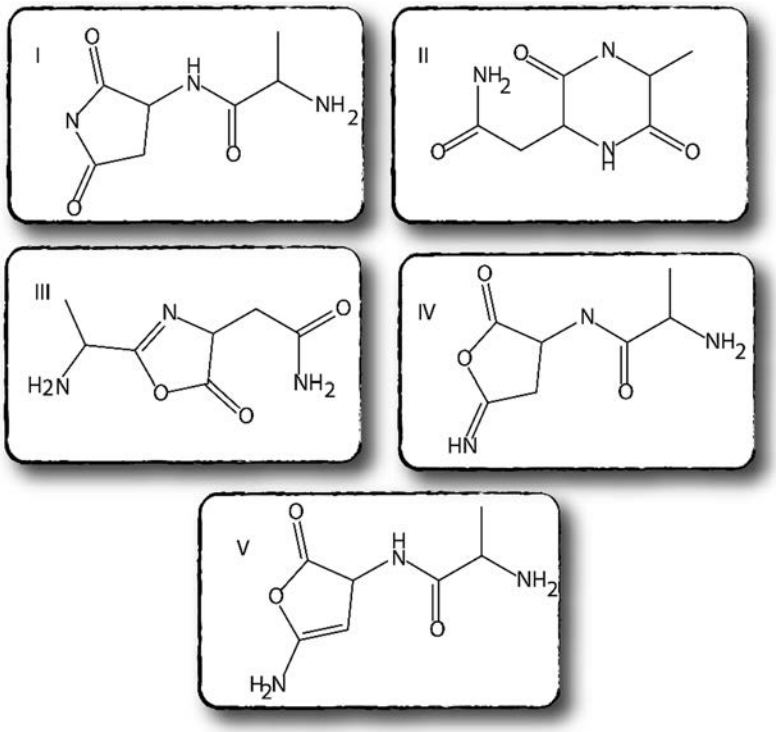


Figure 6.1. Suggested structures and deprotonation sites for b_2 peptide fragments generated from deprotonated ANA: succinimide structure deprotonated at succinimide nitrogen (I), diketopiperazine structure deprotonated at diketopiperazine ring nitrogen (II), oxazolone deprotonated at α-carbon embedded in oxazolone ring (III), imino-γ-butyrolactone structure deprotonated at the peptide bond nitrogen (IV) and amino-γ-butyrolactone deprotonated at the α-carbon in the butyrolactone ring.

In Figure 6.3, the experimental IRMPD spectrum of the m/z 184 anion is compared with calculated vibrational spectra for selected candidate structures: succinimide deprotonated at the succinimide nitrogen atom (Figure 6.3a and 6.3b, showing two different conformations of structure I), diketopiperazine deprotonated at the diketopiperazine nitrogen atom (Figure 6.3c, structure II), oxazolone deprotonated at the α-carbon atom in the oxazolone ring (Figure 6.3d, structure III), imino-γ-butyrolactone deprotonated at the peptide bond nitrogen atom (Figure 6.3e, structure IV) and amino-γ-butyrolactone deprotonated at the α-carbon atom in the butyrolactone ring (Figure 6.3f, structure V). Relative free energy values are given for each of the structures in Figure 6.3. Other possible deprotonation sites for each of the presented isomeric structures were also investigated, but resulted in calculated spectra that clearly did not match with the experiment and that were much higher in relative free energy; these structures are therefore not presented.

Inspection of the six calculated spectra clearly suggests that only the succinimide conformers in panels a and b are in reasonable agreement with the experiment. The two conformers mainly differ in the orientation of the N-terminal amino-group with respect to the rest of the system, changing the H-bonding of the N-terminus. These succinimide structures have the lowest free energies of all isomers and their free energy difference is only 9 kJ/mol. Their calculated spectra are rather similar. Here, we analyze the observed

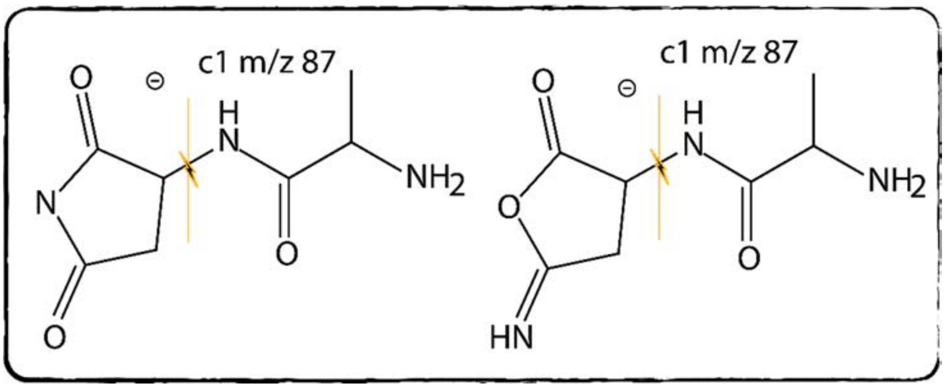


Figure 6.2. Succinimide or γ-butyrolactone? Observed IRMPD fragment ions (c_1 ion and m/z 96) of the b_2 anion can be explained by Asn side-chain involvement in the dissociation reaction of deprotonated ANA. However, the tandem MS data cannot distinguish between the presented structures - succinimide or imino-γ-butyrolactone.

bands based on the calculated vibrational spectrum of the lowest energy structure presented in panel a. The bands at 1715, 1676 and 1613 cm^{-1} in the computed spectrum correspond to the shoulder at 1730 cm^{-1} and the bands at 1680 and 1613 cm^{-1} in the experimental spectrum and are due to different carbonyl stretching modes: the symmetric carbonyl stretching mode of the succinimide ring, the peptide bond carbonyl stretch and the asymmetric stretching of the succinimide carbonyls, respectively. The band with relatively high intensity at 1491 cm^{-1} (1488 cm^{-1} in the experiment) is assigned to the peptide bond NH wagging with some stretching contribution of the adjacent N-C bond. For the conformer presented in panel a, this NH is hydrogen-bonded only with the succinimide oxygen atom. For the succinimide conformer shown in Figure 6.3b, this NH is in addition hydrogen-bonded to the N-terminal nitrogen atom, which results in a shift of the calculated NH wagging frequency to 1509 cm^{-1}, making the match between experiment and theory slightly worse for this conformer. In the 1200-1300 cm^{-1} region, the most prominent band is at 1290 cm^{-1} (1283 cm^{-1} in the experiment) and is due to succinimide NC stretching. Further small features that appear below 1290 cm^{-1} are due to different C-H vibrations coupled to N-H vibrations and they form a broad shoulder around 1230 cm^{-1}. In the computed spectrum, one band appears that does not match our experimental data at 913 cm^{-1}. This band is attributed to the NH_2 bending mode, which is often a vibration with high anharmonic behavior[59,60]. Hence, we conclude that the agreement between theory and experiment is excellent for the lowest energy conformer in Figure 6.3a, particularly in the critically diagnostic spectral region above 1400 cm^{-1}. However, considering the small relative energy difference and the spectral similarity, admixture of the conformer in Figure 6.3b cannot be excluded. The b_2-ion of deprotonated ANA can thus safely be concluded to possess a succinimide structure deprotonated on the succinimide nitrogen atom.

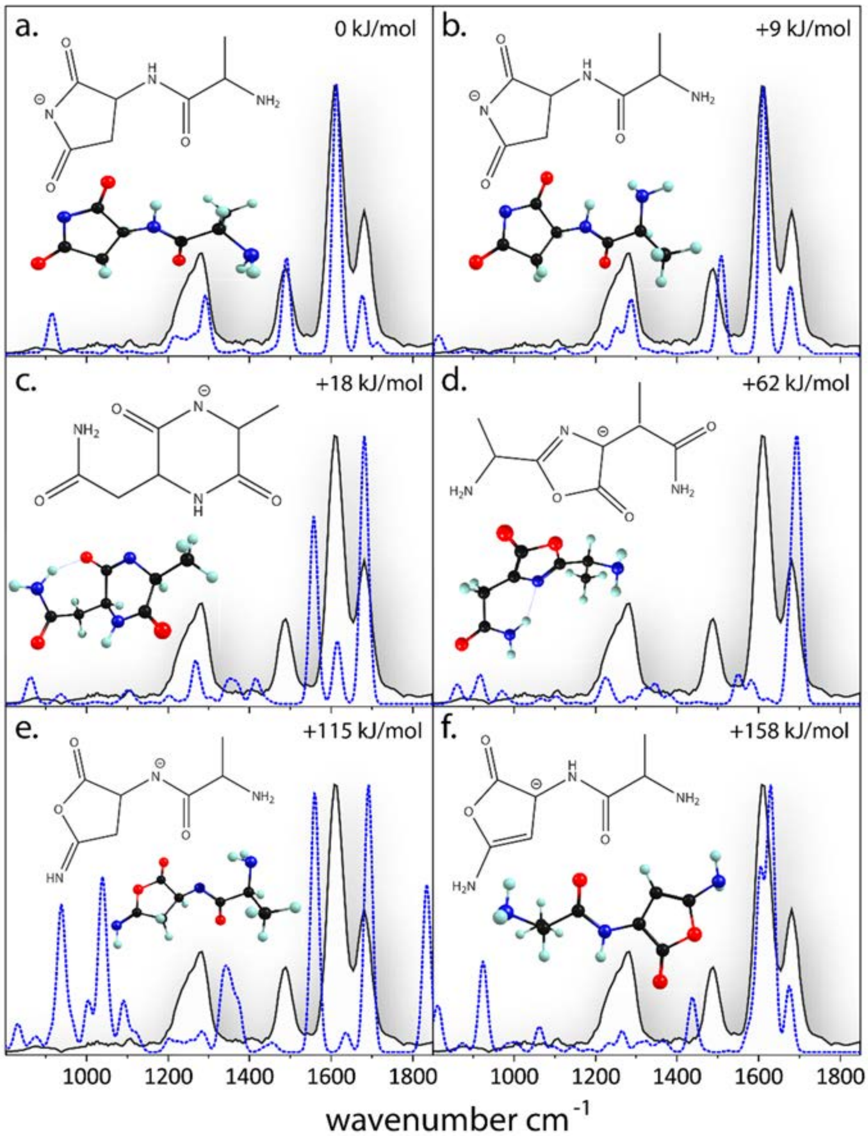
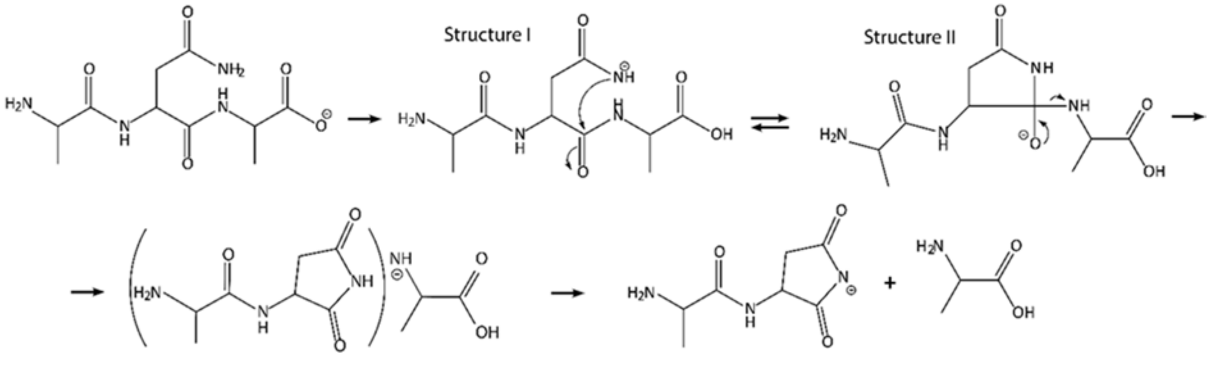


Figure 6.3. IRMPD spectrum of the b_2 ion of deprotonated ANA (black in all panels) compared to DFT-calculated spectra (blue dashed line) for candidate structures. Calculated relative free energies are given for each of the structures.

6.3.2. Proposed mechanism for b_2-ion formation

A succinimide structure was previously proposed by Bowie and co-workers for the anion formed by water loss from deprotonated ValAsn[35]. Succinimide formation was also suggested for non-enzymatic deamidation in Asn-containing proteins and peptides in solution[61]. Based on these previous studies, a proposed mechanism for the formation of the succinimide b_2 anion is depicted in Scheme 6.1. The site of initial deprotonation is taken to be the C-terminus, since the C-terminal proton is most acidic in the parent molecule[62,63]. Support for this assumption is further provided by the spectroscopic observation that the deprotonation site of Ala_3 and AlaTyrAla, the latter containing a moderately acidic Tyr residue, is the C-terminus[36,37].



Scheme 6.1. Possible dissociation mechanism leading to a succinimide b_2-anion structure. The first step is proton transfer from the Asn side-chain nitrogen to the C-terminal carboxylate (structure I). The second step is a nucleophilic attack of the deprotonated amide nitrogen onto the peptide bond carbonyl carbon, which yields a tetrahedral intermediate (structure II), followed by proton transfer, generating the succinimide b_2-anion.

The first step is proton transfer to the C-terminal carboxylate (see Scheme 6.1) and the question is which proton is mobilized. In peptides without acidic residues (Glu, Asp, Tyr), the next most acidic proton is an amide bond proton. In previous spectroscopic experiments on anionic N-terminal a- and c-type peptide fragments, deprotonation at the peptide bond nitrogen has indeed been clearly identified[36,38]. For peptides containing Asn or Gln residues with an amide moiety in their side chain, one may wonder whether the side chain amide is more acidic than the peptide bond amide. The gas-phase acidities for the model compounds acetamide ($CH_3C(=O)NH_2$) and N-methyl acetamide ($CH_3C(=O)NHCH_3$) were found to be identical within 2 kJ/mol (i.e. within experimental accuracy)[64]. It is thus conceivable that upon collisional activation, the side chain amide proton is able to migrate to the C-terminal carboxylate, as sketched in Scheme 6.1. The second step is then a nucleophilic attack of the deprotonated amide nitrogen on the peptide bond carbonyl carbon, which produces a tetrahedral intermediate (Scheme 6.1). This is followed by charge and proton transfer that leads to rearrangement, fragmentation and formation of the succinimide b_2-anion.

It is interesting to note that this mechanism is analogous to that described by the 'mobile proton' theory in the fragmentation of protonated peptides[65-67]. Although deprotonated species lack one proton instead of having one extra, Bache and coworkers showed that hydrogen scrambling occurs also in anionic peptides upon collisional excitation[68].

6.3.3. b_3 anion from AlaAlaAsnAla

Structures that we will consider as candidates for the b_3 fragment from deprotonated AANA are analogous to those considered for the b_2 fragments of ANA, with the exception that a diketopiperazine structure cannot be formed for a b_3 anion. This structure is therefore excluded from the following discussion.

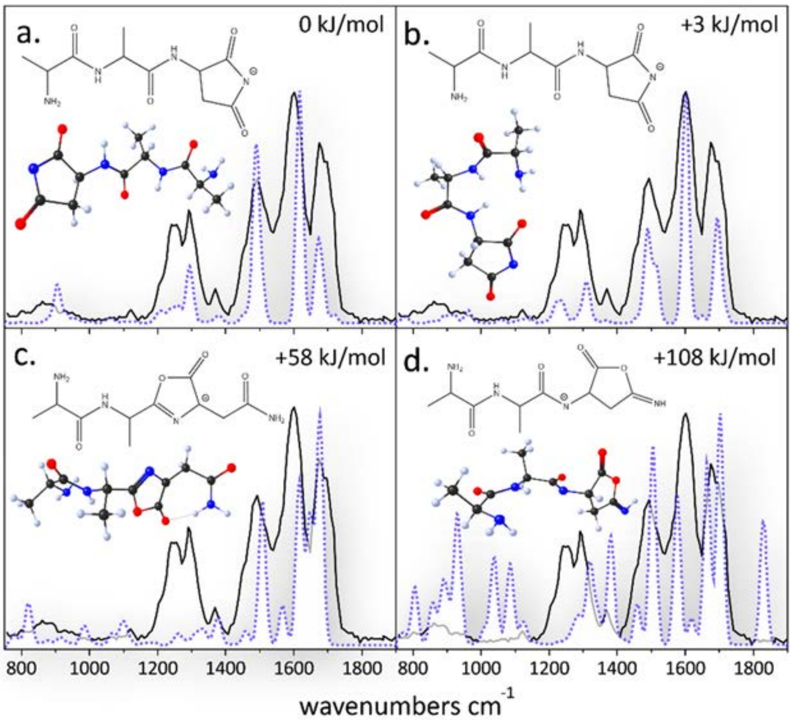


Figure 6.4. IRMPD spectrum of the b_3 ion of deprotonated AANA (black in all panels) compared to DFT-calculated spectra (blue dashed line) for candidate structures. Calculated relative free energies are given for each of the structures.

Upon irradiation with a 0.1s long CO_2 laser pulse, the main dissociation channels observed for deprotonated AANA (m/z 344) were loss of ammonia (m/z 327), the b_3 fragment (m/z 255), the c_2 fragment (m/z 158) and m/z 96. Upon IRMPD of the isolated b_3 fragment, dissociation into the c_2-ion and m/z 96 was observed. This resembles the fragmentation pathway observed for the b_2 anion from ANA (c_1 and 96 m/z), suggestive of similar fragment structures.

Figure 6.4 presents a direct comparison of the experimental IRMPD spectrum of the b_3-anion with computed IR spectra for various candidate structures: succinimide deprotonated at the succinimde nitrogen atom (Figure 6.4a and 6.4b), oxazolone deprotonated at the α-carbon atom in the oxazolone ring (Figure 6.4c) and imino-γ-butyrolactone deprotonated at the peptide bond nitrogen atom (Figure 6.4d). Relative free energy values are given for each of the structures. It is immediately obvious that good agreement between experiment and theory is only achieved for the succinimide structures in panels **a** and **b**. These two conformations of the succinimide isomer also possess the lowest energies, while the computed energy difference between them is insignificantly small (3 kJ/mol). In addition, the two conformers have nearly identical computed spectra. In the diagnostic 1500–1700 cm^{-1} region, it appears that the best match is provided by conformer **b**. However, based on experimental and calculated spectra, it cannot be excluded that a mixture of these conformers is present, which is not unlikely on the basis of their similar stabilities. According to the calculated vibrational frequencies and normal modes, the most intense bands in the experimental spectrum at 1675, 1601 and 1493 cm^{-1} originate from peptide bond carbonyl

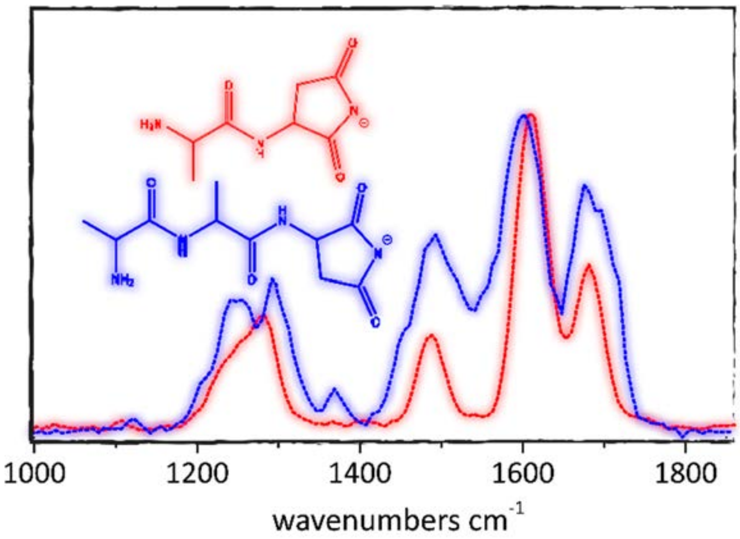


Figure 6.5. Overlaid IRMPD spectra of the b_3 (blue dotted line) from AANA and b_2 ion of deprotonated ANA (red dash-dot line). Consistency of the most intense bands of these two spectra suggests similarity of the corresponding fragment structures.

stretching, asymmetric succinimide carbonyl stretching and NH bending modes, respectively. The next band of interest lies at 1296 cm^{-1} and is due to the succinimide NC stretching vibration. The broader and relatively intense band at about 1240 cm^{-1} is due to two unresolved vibrational bands, which are identified as succinimide C-H rocking and N-H wagging. Less intense bands between 900-1200 cm^{-1} show reasonable overlap with the computed spectra and are a result of different C-H and N-H coupled vibrations.

Figure 6.5 shows an overlay of the experimental spectra of b_2 from ANA and b_3 from AANA. Consistency of the most intense bands in these two spectra clearly suggests a large extent of similarity between the isomeric structures of the b_2- and b_3-fragments. Slight offsets between the bands is due to the additional amide moiety in the b_3 ion as compared to b_2, the higher level of conformational freedom of the b_3 anion as well as the possibility to establish additional hydrogen bonds, which result in shifts of particular absorptions of the groups involved.

6.3.4. Spectroscopic identification of the m/z 96 fragment

Further experimental confirmation of our assignment as succinimide structures for b_2 and b_3 fragment anions can be obtained by establishing the structure of the m/z 96 fragment that they appear to have in common. To this end, the experimental spectrum of the m/z 96 fragment generated from the b_3 anion of deprotonated AANA was recorded and is presented in Figure 6.6. Isolation of the m/z 96 fragment and subsequent resonant IRMPD with FELIX generates a single fragment anion at m/z 42, corresponding to NCO^-. The IRMPD spectrum was produced by plotting the wavelength-dependent yield of this fragment.

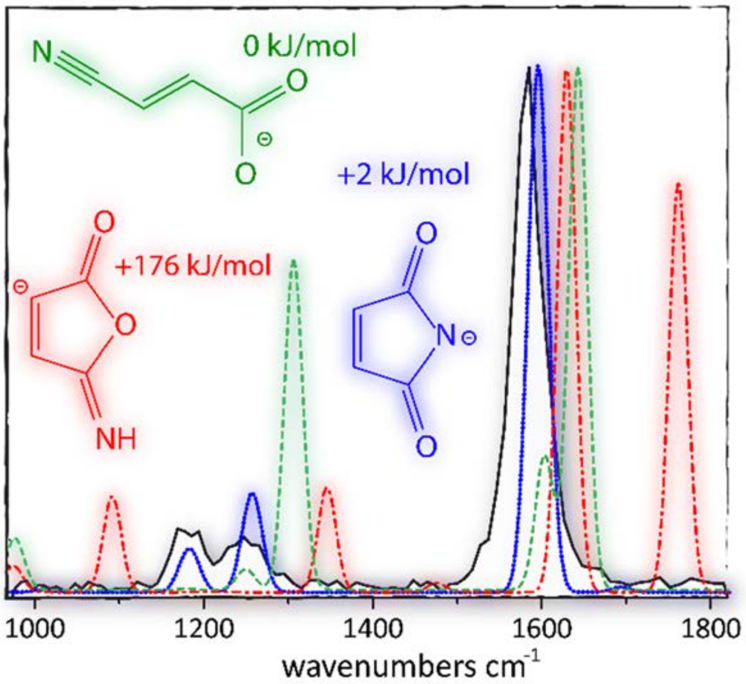
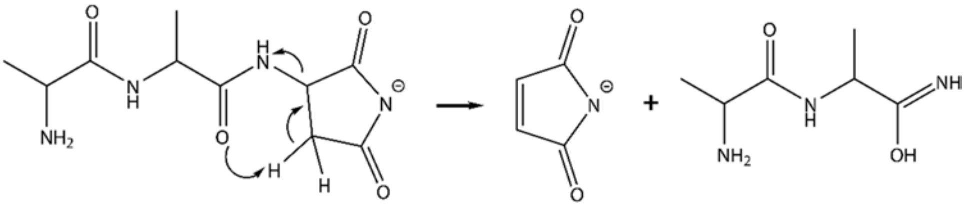


Figure 6.6. IRMPD spectrum of the m/z 96 anion of deprotonated AANA (black) compared to DFT-calculated IR spectra for candidate structures: maleimide deprotonated at the nitrogen atom (blue dotted), 5-imino-furan-2-one deprotonated at the carbon atom adjacent to the carbonyl group (red dot-dashed) and a possible linear carboxylate structure (green dashed).



Scheme 6.2. Suggested fragmentation pathway leading to the formation of the maleimide anion from the b$_3$ fragment of deprotonated AANA.

From the IRMPD spectrum recorded and its comparison to computed spectra, it is unambiguous that this ion is the maleimide (3,4-dehydrogenated succinimide) anion deprotonated at the nitrogen atom, and not the alternative 5-imino-furan-2-on anion that would have been expected from a γ-butyrolactone *b*-fragment structure. Isomeric structures for m/z 96, also including an alternative linear carboxylate structure, are presented in Figure 6.6 and their corresponding computed vibrational spectra are marked by color. Only three bands are observed in the experimental spectrum, positioned at 1185, 1247 and 1585 cm^{-1}. We attribute them to vibrations of the maleimide anion as the C-N-C asymmetric stretching mode with some contribution from the C-H rocking motion, the C-N-C symmetric stretching mode and the C=O asymmetric stretching mode, respectively. A straightforward mechanism can be proposed to explain the formation of the m/z 96 fragment anion from the succinimide b$_3$ peptide fragment anion (see Scheme 6.2).

6.4. Conclusion

IRMPD spectroscopy using a free electron laser was applied to investigate the structures of short anionic *b*-type peptide fragments generated from peptides containing Asn, in particular from ANA and AANA. Comparison with calculated spectra show unambiguously that the b_2 fragment from ANA and the b_3 fragment from AANA do not possess the classical oxazolone structure, but instead adopt cyclic succinimide structures, deprotonated on the succinimide nitrogen. These findings are consistent with tandem MS data of the b_2 and b_3 fragments, i.e. with formation of c_1 (from b_2) and c_2 (from b_3) anions, as well as with maleimide formation at m/z 96 (from both b_2 and b_3).

The succinimide structures are likely formed via nucleophilic attack of the side chain amide nitrogen, which becomes deprotonated upon collsional activation and proton migration. For both peptide fragment anions, the thermochemically most favored fragment is formed. Deprotonation at the succinimide ring is favored over other deprotonation sites (amide proton of the peptide linkage in particular) since it provides delocalization of the negative charge over five atoms (succinimide nitrogen and two carbonyl groups), which results in enhanced gas-phase stability. A detailed reaction pathway would require further computational investigation of the potential energy surfaces involved in the fragmentation reactions, which is beyond the scope of this study.

It is of interest to investigate fragment ion structures of deprotonated peptides containing other functionalized residues. In light of the present findings for Asn-containing peptides, the behavior of the glutamine (Gln) residue is of particular interest as it also contains an amide moiety, though linked to the backbone by a longer alkyl chain. For protonated peptides containing Asn and Gln, similar involvement of the side-chain amide in the reactions forming *b*-ions was recently established using ion spectroscopy[14]. The role of Gln in the dissociation of deprotonated peptides is currently under study in our group.

Acknowledgments

We gratefully thank the FELIX staff, in particular Drs. A.F.G. van der Meer, B. Redlich and G. Berden for their expert technical support. JO acknowledges the Stichting Physica for support. Financial support for this project has been provided by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) in the form of VICI grant no. 724.011.002. This work is part of the research program of FOM, which is financially supported by NWO

REFERENCES

(1) Aebersold, R.; Goodlett, D. R. *Mass spectrometry in proteomics* Chem. Rev. **2001**, *101*, 269.

(2) Steen, H.; Mann, M. *The ABC's (and XYZ's) of peptide sequencing* Nat. Rev. Mol. Cell Biol. **2004**, *5*, 699.

(3) Yates, J. R. *Mass spectrometry and the age of the proteome* J. Mass Spectrom. **1998**, *33*, 1.

(4) Barlow, C. K.; O'Hair, R. A. J. *Gas-Phase Peptide Fragmentation: How Understanding the Fundamentals Provides a Springboard to Develping New Chemistry and Novel Proteomics Tools* J. Mass Spectrom. **2008**, *43*, 1301.

(5) Eyler, J. R. *Infrared multiple photon dissociation spectroscopy of ions in Penning traps* Mass Spectrom. Rev. **2009**, *28*, 448.

(6) Polfer, N. C.; Oomens, J. *Vibrational spectroscopy of bare and solvated ionic complexes of biological relevance* Mass Spectrom. Rev. **2009**, *28*, 468.

(7) Polfer, N. C.; Oomens, J. *Reaction products in mass spectrometry elucidated with infrared spectroscopy* Phys. Chem. Chem. Phys. **2007**, *9*, 3804.

(8) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *Spectroscopic and theoretical evidence for oxazolone ring formation in collision-induced dissociation of peptides* J. Am. Chem. Soc. **2005**, *127*, 17154.

(9) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *Infrared spectroscopy and theoretical studies on gas-phase protonated Leu-enkephalin and its fragments: Direct experimental evidence for the mobile proton* J. Am. Chem. Soc. **2007**, *129*, 5887.

(10) Roepstorff P; J., F. *Proposal for a common nomenclature for sequence ions in mass spectra of peptides* Biomed. Mass Spectrom. **1984**, *11*, 601.

(11) Oomens, J.; Young, S.; Molesworth, S.; van Stipdonk, M. *Spectroscopic evidence for an oxazolone structure of the b_2 fragment ion from protonated tri-alanine* J. Am. Soc. Mass Spectrom. **2009**, *20*, 334.

(12) Yoon, S. H.; Chamot-Rooke, J.; Perkins, B. R.; Hilderbrand, A. E.; Poutsma, J. C.; Wysocki, V. H. *IRMPD spectroscopy shows that AGG forms an oxazolone b_2^+ ion* J. Am. Chem. Soc. **2008**, *130*, 17644.

(13) Perkins, B. R.; Chamot-Rooke, J.; Yoon, S. H.; Gucinski, A. C.; Somogyi, A.; Wysocki, V. H. *Evidence of diketopiperazine and oxazolone structures for HA b_2^+ ion* J. Am. Chem. Soc. **2009**, *131*, 17528.

(14) Grzetic, J.; Oomens, J. *Spectroscopic Identification of Cyclic Imide b_2-Ions from Peptides Containing Gln and Asn Residues* J. Am. Soc. Mass Spectrom **2013**, *24*, 1228.

(15) Harisson, A. G. *Peptide Sequence Scrambling Through Cyclization of b_5 Ions* J. Am. Soc. Mass Spectrom. **2008**, *19*, 1776.

(16) Bleiholder, C.; Osburn, S.; Williams, T. D.; Suhai, S.; Van Stipdonk, M.; Harrison, A. G.; Paizs, B. *Sequence-scrambling fragmentation pathways of protonated peptides* J. Am. Chem. Soc. **2008**, *130*, 17774.

(17) Molesworth, S.; Leavitt, C. M.; Groenewold, G. S.; Oomens, J.; Steill, J. D.; Van Stipdonk, M. *Spectroscopic evidence for mobilization of amide position protons during CID of model peptide ions* J. Am. Soc. Mass Spectrom. **2009**, *20*, 1841.

(18) Chen, X.; Steill, J. D.; Oomens, J.; Polfer, N. C. *Oxazolone versus macrocycle structures for Leu-Enkephalin b_2-b_4: Insights from infrared multiple-photon dissociation spectroscopy and gas-phase hydrogen/deuterium exchange* J. Am. Soc. Mass Spectrom. **2010**, *21*, 1313.

(19) Erlekam, U.; Bythell, B. J.; Scuderi, D.; Van Stipdonk, M.; Paizs, B.; Maitre, P. *Infrared spectroscopy of fragments of protonated peptides: Direct evidence for macrocyclic structures of* b_5 *ions* J. Am. Chem. Soc. **2009**, *131*, 11503.

(20) Verkerk, U. H.; Siu, C. K.; Steill, J. D.; El Aribi, H.; Zhao, J.; Rodriguez, C. F.; Oomens, J.; Hopkinson, A. C.; Siu, K. W. M. a_2 *Ion Derived from Triglycine: An* N_1*-Protonated 4-Imidazolidinone* J. Phys. Chem. Lett. **2010**, *1*, 868.

(21) Bythell, B. J.; Maitre, P.; Paizs, B. *Cyclization and Rearrangement Reactions of* a_n *Fragment Ions of Protonated Peptides* J. Am. Chem. Soc. **2010**, *132*, 14766.

(22) Andreazza, H. J.; Wang, T.; Bilusich, D.; Hoffmann, P.; Bowie, J. H. *Negative ion fragmentations of deprotonated peptides containing post-translational modifications: diphosphorylated systems containing Ser, Thr and Tyr. A characteristic phosphate/phosphate cyclisation. A joint experimental and theoretical study* Rapid Commun. Mass Spectrom. **2009**, *23*, 1825.

(23) Brinkworth, C. S.; Dua, S.; McAnoy, A. M.; Bowie, J. H. *Negative ion fragmentations of deprotonated peptides: backbone cleavages directed through both Asp and Glu* Rapid Commun. Mass Spectrom. **2001**, *15*, 1965.

(24) Bowie, J. H.; Brinkworth, C. S.; Dua, S. *Collision-induced fragmentations of the* $(M\text{-}H)^-$ *parent anions of underivatized peptides: An aid to structure determination and some unusual negative ion cleavages* Mass Spectrom. Rev. **2002**, *21*, 87.

(25) Pu, D.; Clipston, N. L.; Cassady, C. J. *A comparison of positive and negative ion collision-induced dissociation for model heptapeptides with one basic residue* J. Mass Spectrom. **2010**, *45*, 297.

(26) Song, H.; Håkansson, K. *Electron detachment dissociation and negative ion infrared multiphoton dissociation of electrosprayed intact proteins* Anal. Chem. **2011**, *84*, 871.

(27) Harrison, A. *Effect of phenylalanine on the fragmentation of deprotonated peptides* J. Am. Soc. Mass. Spectrom. **2002**, *13*, 1242.

(28) Harrison, A. G.; Young, A. B. *Fragmentation reactions of deprotonated peptides containing aspartic acid* Int. J. Mass Spectrom. **2006**, *255-256*, 111.

(29) Harrison, A. G. *Sequence-specific fragmentation of deprotonated peptides containing H or alkyl side chains* J. Am. Soc. Mass Spectrom. **2001**, *12*, 1.

(30) Li, Z.; Yalcin, T.; Cassady, C. J. *C-terminal amino acid residue loss for deprotonated peptide ions containing glutamic acid, aspartic acid, or serine residues at the C-terminus* J. Mass Spectrom. **2006**, *41*, 939.

(31) Men, L.; Wang, Y. *Fragmentation of the deprotonated ions of peptides containing cysteine, cysteine sulfinic acid, cysteine sulfonic acid, aspartic acid, and glutamic acid* Rapid Commun. Mass Spectrom. **2006**, *20*, 777.

(32) Pu, D.; Cassady, C. J. *Negative ion dissociation of peptides containing hydroxyl side chains* Rapid Commun. Mass Spectrom. **2008**, *22*, 91.

(33) Waugh, R. J.; Bowie, J. H.; Hayes, R. N. *Collision induced dissociations of deprotonated peptides: Dipeptides containing phenylalanine, tyrosine, histidine and tryptophan* Int. J. Mass Spectrom. **1991**, *107*, 333.

(34) Harrison, A. G.; Siu, K. W. M.; El Aribi, H. *Amide bond cleavage in deprotonated tripeptides: a newly discovered pathway to " $b_{(2)}$ ions* Rapid Comm. Mass Spectrom. **2003**, *17*, 869.

(35) Waugh, R. J.; Bowie, J. H.; Gross, M. L. *Collision-Induced Dissociations of Deprotonated Peptides. Dipeptides Containing Asn, Arg and Lys* Aust. J. Chem. **1993**, *46*, 693.

(36) Oomens, J.; Steill, J. D. *The structure of deprotonated tri-alanine and its* a_3 *fragment anion by IR spectroscopy* J. Am. Soc. Mass Spectrom. **2010**, *21*, 698.

(37) Grzetic, J.; Oomens, J. *Spectroscopic evidence for an oxazolone structure in anionic b-type peptide fragments* J. Am. Soc. Mass Spectrom. **2012**, *23*, 290.

(38) Grzetic, J.; Oomens, J. *Structure of anionic c-type peptide fragments elucidated by IRMPD spectroscopy* Int. J. of Mass Spectrom. **2012**, *316-318*, 216.

(39) Farrugia, J. M.; OHair, R. A. J.; Reid, G. E. *Do all* b_2 *ions have oxazolone structures? Multistage mass spectrometry and ab initio studies on protonated N-acyl amino acid methyl ester model systems* Int. J. Mass Spectrom. **2001**, *210-211*, 71.

(40) Harrison, A. G. *Fragmentation reactions of protonated peptides containing glutamine or glutamic acid* J. Mass Spectrom. **2003**, *38*, 174.

(41) Lee, Y. J.; Lee, Y. M. *Formation of* c_1 *fragment ions in collision-induced dissociation of glutamine-containing peptide ions: a tip for de novo sequencing* Rapid Comm. Mass Spectrom. **2004**, *18*, 2069.

(42) Winter, D.; Seidler, J.; Hahn, B.; Lehmann, W. *Structural and mechanistic information on* c_1 *ion formation in collision-induced fragmentation of peptides* J. Am. Soc. Mass Spectrom. **2010**, *21*, 1814.

(43) Robinson, N. E.; Robinson, A. B. *Molecular Clocks: Deamidation of Asparaginyland Glutaminyl Residues in Peptides and Proteins, Althouse Press, Cave Junction, OR.* **2004**.

(44) Dunkelberger, E. B.; Buchanan, L. E.; Marek, P.; Cao, P.; Raleigh, D. P.; Zanni, M. T. *Deamidation Accelerates Amyloid Formation and Alters Amylin Fiber Structure* Journal of the American Chemical Society **2012**, *134*, 12658.

(45) Valle, J. J.; Eyler, J. R.; Oomens, J.; Moore, D. T.; van der Meer, A. F. G.; von Helden, G.; Meijer, G.; Hendrickson, C. L.; Marshall, A. G.; Blakney, G. T. *Free electron laser-Fourier transform ion cyclotron resonance mass spectrometry facility for obtaining infrared multiphoton dissociation spectra of gaseous ions* Rev. Sci. Instrum. **2005**, *76*, 023103.

(46) Oepts, D.; van der Meer, A. F. G.; van Amersfoort, P. W. *The infrared free-electron laser facility FELIX* Infrared Phys. Technol. **1995**, *36*, 297.

(47) Marshall, A. G.; Wang, T. C. L.; Ricca, T. L. *Tailored excitation for Fourier transform ion cyclotron mass spectrometry* J. Am. Chem. Soc. **1985**, *107*, 7893.

(48) Shaffer, C. J.; Revesz, A.; Schröder, D.; Severa, L.; Teply, L.; Zins, E.-L.; Jasikova, L.; Roithova, J. *Can Hindered Intramolecular Vibrational Energy Redistribution Lead to Non-Ergodic Behavior of Medium-Sized Ion Pairs?* Angew. Chem. Int. Ed. **2012**, *51*, 10050.

(49) Mize, T. H.; Taban, I.; Duursma, M.; Seynen, M.; Konijnenburg, M.; Vijftigschild, A.; Doornik, C. V.; Rooij, G. V.; Heeren, R. M. A. *A modular data and control system to improve sensitivity, selectivity, speed of analysis, ease of use, and transient duration in an external source FTICR-MS* Int. J. Mass Spectrom. **2004**, *235*, 243.

(50) M. J. Frisch, G. W. T., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, *Gaussian 09, Revision A.1* **2009**.

(51) Merrill, G. N.; Kass, S. R. *Calculated Gas-Phase Acidities Using Density Functional Theory: Is It Reliable?* J. Phys. Chem. **1996**, *100*, 17465.

(52) Oomens, J.; Steill, J. D. *Free carboxylate stretching modes* J. Phys. Chem. A **2008**, *112*, 3281.

(53) Kulik, W.; Heerma, W. *Fast Atom Bombardment Tandem Mass-Spectrometry for Amino-Acid Sequence Determination in Tripeptides* Biomed. Environ. Mass Spectrom. **1989**, *18*, 910.

(54) Waugh, R. J.; Bowie, J. H. *A review of the collision induced dissociations of deprotonated dipeptides and tripeptides. An aid to structure determination.* Rapid Commun. Mass Spectrom. **1994**, *8*, 169.

(55) Bowie, J. H.; Brinkworth, C. S.; Dua, S. *Collision-induced fragmentations of the (M - H)- parent anions of underivatized peptides: an aid to structure determination and some unusual negative ion cleavages* Mass Spectrom. Rev. **2002**, *21*, 87.

(56) Eckersley, M.; Bowie, J. H.; Hayes, R. N. *Collision-Induced Dissociations of Deprotonated Peptides, Dipeptides and Tripeptides with Hydrogen and Alkyl a-Groups-an Aid to Structure Determination* Org. Mass Spectrom. **1989**, *24*, 597.

(57) Harrison, A.; Young, A. *Fragmentation of deprotonated N-benzoylpeptides: Formation of deprotonated oxazolones* J. Am. Soc. Mass. Spectrom. **2004**, *15*, 446.

(58) Wabnitz, P. A.; Waugh, R. J.; Eckersley, M. A.; Dua, S.; Blumenthal, T.; Bowie, J. H. *The negative ion mass spectra of deprotonated 2,5-diketopiperazines* Int. J. Mass Spectrom. Ion Processes **1996**, *154*, 193.

(59) Sinclair, W. E.; Pratt, D. W. J. Chem. Phys. **1996**, *105*, 7942.

(60) Piest, H.; von Helden, G.; Meijer, G. *Infrared spectroscopy of jet-cooled neutral and ionized aniline–Ar* J. Chem. Phys. **1999.**, *110*, 2010.

(61) Capasso, S.; Mazzarella, L.; Sica, F.; Zagari, A.; Salvadori, S. *Kinetics and mechanism of succinimide ring formation in the deamidation process of asparagine residues* J. Chem. Soc., Perkin Trans. 2 **1993**, 679.

(62) O'Hair, R. A. J.; Bowie, J. H.; Gronert, S. *Gas phase acidities of the alpha amino acids* Int. J. Mass Spectrom. Ion Proc. **1992**, *117*, 23.

(63) Jones, C. M.; Bernier, M.; Carson, E.; Colyer, K. E.; Metz, R.; Pawlow, A.; Wischow, E. D.; Webb, I.; Andriole, E. J.; Poutsma, J. C. *Gas-phase acidities of the 20 protein amino acids* Int. J. Mass Spectrom. **2007**, *267*, 54.

(64) Decouzon, M.; Exner, O.; Gal, J.-F.; Maria, P.-C. *The Gas-Phase Acidity and the Acidic Site of Acetohydroxamic Acid: A FT-ICR Study* J. Org. Chem. **1990**, *55*, 3980.

(65) Burlet, O.; Orkiszewski, R. S.; Ballard, K. D.; Gaskell, S. *Charge promotion of low-energy fragmentations of peptide ions* Rapid Commun. Mass Spectrom. **1992**, *6*, 658.

(66) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *Mobile and localized protons: a framework for understanding peptide dissociation* Journal of Mass Spectrometry **2000**, *35*, 1399.

(67) Boyd, R.; Somogyi, A. *The mobile proton hypothesis in fragmentation of protonated peptides: a perspective* J. Am. Soc. Mass Spectrom. **2010**, *21*, 1275.

(68) Bache, N.; Rand, K. D.; Roepstorff, P.; Ploug, M.; Jørgensen, T. J. D. *Hydrogen Atom Scrambling in Selectively Labeled Anionic Peptides Upon Collisional Activation by MALDI Tandem Time-of-Flight Mass Spectrometry* J. Am. Soc. Mass. Spectrom. **2008**, *19*, 1719.

Chapter 7.

Spectroscopic identification of cyclic imide b_2-ions from peptides containing Gln and Asn residues

In mass-spectrometry based peptide sequencing, formation of b- and y-type fragments by cleavage of the amide C-N bond constitutes the main dissociation pathway of protonated peptides under low-energy collision induced dissociation (CID). The structure of the b_2 fragment ion from peptides containing glutamine (Gln) and asparagine (Asn) residues is investigated here by infrared ion spectroscopy using the free electron laser FELIX. The spectra are compared with theoretical spectra calculated using density functional theory for different possible isomeric structures as well as to experimental spectra of synthesized model systems. The spectra unambiguously show that the b_2-ions do not possess the common oxazolone structure, nor do they possess the alternative diketopiperzazine structure. Instead, cyclic imide structures are formed through nucleophilic attack by the amide nitrogen atom of the Gln and Asn side chains. The alternative pathway involving nucleophilic attack from the side-chain amide oxygen atom leading to cyclic isoimide structures, which had been suggested by several authors, can clearly be excluded based on the present IR spectra. This mechanism is perhaps surprising as the amide oxygen atom is considered to be the better nucleophile; however, computations show that the products formed via attack by the amide nitrogen are considerably lower in energy. Hence, b_2-ions with Asn or Gln in the second position form structures with a five-membered succinimide or a six-membered glutarimide ring, respectively. b_2-ions formed from peptides with Asn in the first position are spectroscopically shown to possess the classical oxazolone structure. *

---

*Adapted from: J.Grzetic and J.Oomens, *Spectroscopic identification of cyclic imide b_2-ions from peptides containing Gln and Asn residues,* J. Am. Soc. Mass Specctrom. (2013), 24, 1228-1241

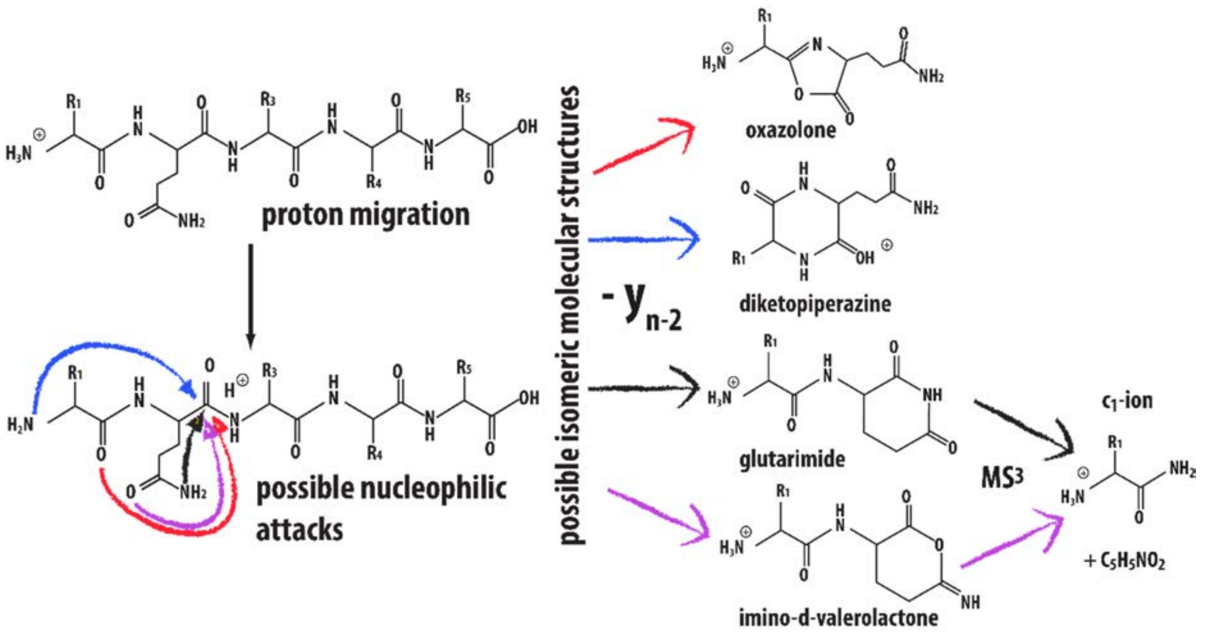
7.1. Introduction

Current proteomics research relies heavily on mass spectrometry (MS) based peptide sequencing methods.[2,3] Practical applications make use of the differences in the m/z values of peptide fragments generated by collision-, electron-, surface- or photon-induced dissociation of protonated peptides. In most applications, protein identification is then effected by database matching. Along with the practical application of MS-based peptide sequencing, much effort, both experimental and theoretical, has been devoted to a fundamental understanding of the dissociation reactions occurring in the mass spectrometer. Such experimental and computational studies generally address the energetics and dynamics of the dissociation reactions including the minimum and transition-state structures along the reaction path.[4-13]

Under low-energy collision conditions, protonated peptides fragment predominantly by cleavage of the backbone amide bond generating so-called b- and y-type fragments,[3] corresponding to the charged N- and C-terminal products, respectively, of the dissociation reaction. The mobile proton model[4,5,14] predicts that upon collisional activation, the added proton migrates from the most basic site in the peptide to one of the amide linkages (either the amide nitrogen or oxygen atom), weakening the amide bond and making the amide carbonyl C-atom a strong electrophile (see Scheme 7.1).[7,13] This carbonyl C-atom becomes subject of nucleophilic attack leading to the reactive configuration. The oxygen atom of the adjacent amide carbonyl is generally assumed to act as the nucleophile in this process, leading to the formation of an oxazolone, five-membered ring structure as shown in Scheme 7.1.[11,15,16] However, the peptide generally contains several other nucleophiles that could possibly attack the electrophilic carbon atom as well, leading to isomeric, non-classical, b-type fragments. Thus, while y-type fragments are generally assumed to possess the structure of a truncated peptide, the structure of the N-terminal b-fragment has been a subject of much debate.[11,15,17]

When unprotonated, the N-terminal amine nitrogen atom is a strong nucleophile. For oxazolone b-type fragments containing more than 4 or 5 residues, this nitrogen atom has been suggested to attack the carbonyl C-atom of the oxazolone ring forming a cyclic peptide.[18,19] Experimental evidence for the formation of these cyclic structures, a potential source of the non-sequence ions, has been provided by multistage CID MS,[18-21] H/D exchange,[22,23] electron capture dissociation (ECD)[24] and IR spectroscopic[22,23,25-27] measurements.

The structure of shorter b-type fragments, in particular the b_2 fragment ion, has also been under much debate.[17] A statistical analysis of a large set of tryptic peptide tandem MS spectra revealed that the occurrence of this fragment ion is anomalously high as compared to longer b-fragments, which was suggested to be due to bifurcating reaction pathways



Scheme 7.1. Possible dissociation/rearrangement reactions based on the 'mobile proton' model, leading to different b_2-ion structures. In addition to the classical oxazolone and diketopiperazine structures, a Gln residue in the second position allows for the formation of glutarimide (cyclic imide) and imino-γ-valerolactone (cyclic isoimide) structures. Observation of the c_1 ion in MS^3 experiments can be interpreted as evidence for Gln side-chain induced cyclization reactions, although it cannot distinguish between the imide and isoimide structures.

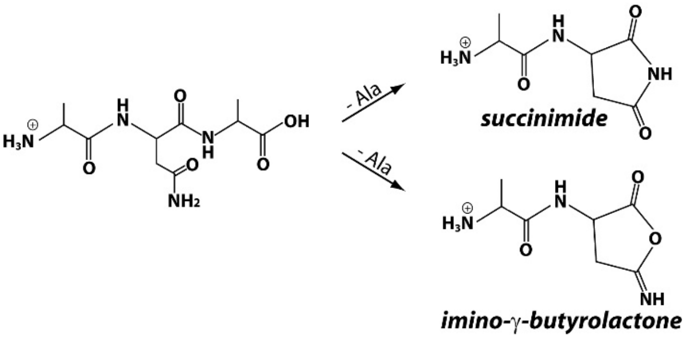
partly leading to a different molecular structure.[28] In particular, the formation of a cyclic dipeptide, i.e. a diketopiperazine structure was suggested. Analogous to the formation of macrocyclic structures, diketopiperazines form when the N-terminal nitrogen atom, instead of the amide carbonyl O-atom, acts as the nucleophile attacking the electrophilic amide carbonyl C-atom (see Scheme 7.1). It is well-known from theoretical studies that the diketopiperazine structure is in fact lower in energy than the oxazolone motif, although its formation is also generally assumed to be kinetically disfavored.[7,16,29-31]

IR photo-dissociation spectroscopy[32-34] of mass-isolated peptide fragment ions has proven to be a diagnostic tool to establish the molecular structures of these gas-phase reaction products.[35,36] Isomeric oxazolone and diketopiperazine b_2 structures are easily distinguished based on their IR spectra. The first b_2-ions investigated with IR spectroscopy contained only Ala and Gly residues and were all found to possess oxazolone structures.[22,37-39] Moreover, ion spectroscopy identified oxazolone b_2-ion structures to be formed from three doubly charged tryptic peptides[40,41] as well as from non-oxazolone b_3 ions.[42] Spectroscopic evidence for a diketopiperazine structure was first found for the b_2 ion of protonated His(Ala)$_4$.[43] It was suggested that since the added proton in the precursor peptide resides on the His side chain, the N-atom of the N-terminal amino group has greater freedom to carry out the nucleophilic attack leading to the diketopiperazine product ion. This scenario is supported by recent spectroscopic studies on b_2-ions containing Arg, which are also predominantly diketopiperazines.[44] Further support is provided by a study on b_2-ions containing histidine analogues, which suggest that the protonated side chain may aid in the trans-to-cis isomerization required for diketopiperazine formation.[45]

Thus, while the side-chain identity has been shown to influence product ion structures indirectly by sequestering the proton,[43,44] virtually no spectroscopic evidence has thus far been found for direct involvement of the side chain in the rearrangement reaction. As the rearrangement reactions are largely driven by nucleophilic attack,[46] side chains of high nucleophilicity yet relatively poor proton affinity are expected to be most prone to become directly involved in the rearrangement reaction. In this study, we therefore focus on glutamine (Gln) and asparagine (Asn) residues, which feature an amide group in their side chain. The amide nitrogen and oxygen atoms form strong nucleophiles, while the proton affinity of the amide group is expected to be lower than that of the N-terminal amine of the peptide.

In their computational survey of possible alternative b_2-ion structures, O'Hair and coworkers indeed suggest cyclization rearrangements induced by the side-chain amide oxygen or nitrogen atom of Gln and Asn residues.[17] Structures formed by such amide N- or O-atom induced nucleophilic attack are however higher in energy than the isomeric diketopiperazine structure, although the energetic ordering of N- and O-atom induced cyclization remains somewhat unclear.[47] Cyclization reactions involving the Gln side-chain N- or O-atom were suggested to occur in analogy to biological degradation of Gln-containing peptides.[48] Further evidence for such processes was found in multistage tandem MS CID experiments of b_2-ions with Gln in the second position, where the unusual formation of a c_1-ion was observed,[49-52] which is of interest in de novo peptide sequencing strategies.[52] Mechanistically, formation of the c_1-ion was interpreted as evidence for the b_2-ion having a non-oxazolone structure; in particular, it was suggested that the b_2-ions having Gln as the second residue possess an imino-δ-valerolactone structure (i.e. a cyclic isoimide, see Scheme 7.1), formed via nucleophilic attack from the side-chain amide oxygen atom.[49,50] However, nucleophilic attack from the Gln side-chain nitrogen atom forming an isomeric glutarimide structure (i.e. a cyclic imide, see Scheme 7.1) is evenly consistent with the observation of the c_1-ion in the MS^3 spectrum and can thus not be excluded.[48] Studies on longer b–type ions containing Gln have encountered the same isomeric question, which remained unresolvable based on MS^n data.[21]

For peptides containing Asn, analogous side-chain induced reactions could lead to non-oxazolone structures, in particular to species containing a five-membered cyclic isoimide (imino-γ-butyrolactone) or imide (succinimide) structure (Scheme 7.2), resulting from side-chain amide oxygen or nitrogen attack, respectively. While both structures were indeed proposed in Ref. [17], MS^3 data reported on Asn containing peptides is ambiguous as to the formation of c-type ions.[50,51] Energetics of imide and isoimide structures relative to the more common oxazolone and diketopiperazine motifs are significantly different for b_2-ions produced from Asn- and Gln-containing peptides.[50]



Scheme 7.2. b_2-ion formation for protonated AlaAsnAla involving the Asn side chain, leading to structures with a five-membered succinimide (cyclic imide) and imino-γ-butyrolactone (cyclic isoimide) ring

Here, we present IR spectra of b_2-ions containing Gln and Asn residues. The b_2-ion structures are determined by comparison of the experimental spectra with spectra calculated for candidate structures as well as with an experimental spectrum of a synthesized model compound. These data suggest that it is indeed the amide side chain that acts as the nucleophile, although not through its oxygen atom as previously suggested, but through its nitrogen atom forming cyclic imide structures: glutarimide for Gln and succinimide for Asn.

7.2. Experimental and computational methods

7.2.1. IRMPD Spectroscopy

Infrared spectra of precursor and fragment ions were obtained via IR multiple photon dissociation (IRMPD) spectroscopy using the Fourier transform ion cyclotron resonance (FTICR) mass spectrometer coupled to the beamline of the infrared free electron laser FELIX.[53,54] Protonated ions were generated by electrospray ionization (ESI) using a Micromass Z-Spray source and peptide solutions of approximately 0.5 mM in acetonitrile/water with about 0.2% acetic acid added. This study focuses on the Gln and Asn containing peptides PheGlnAla, AlaAsnAla and AsnAlaAla possessing an amide moiety in their side chains. The sample peptides were obtained from GeneCust (Dudelange, Luxemburg) and used without further purification. (S)-2-amino-N-((S)-2,6-dioxopiperidin-3-yl)-3-phenylpropanamide, which contains a glutarimide ring and serves as a model for the b_2 fragment of PheGlnAla (vide infra), was synthesized at the University of Amsterdam following the method reported by Fox et al.[55]

The experimental procedure is described in detail elsewhere.[56] Briefly, after generation by ESI, the protonated peptides are accumulated in a linear hexapole trap before being injected into the home-built FTICR mass spectrometer,[54] equipped with a 4.7 T actively-shielded superconducting magnet. In the ICR cell, the ion of interest is mass isolated by a stored waveform inverse Fourier transform (SWIFT) excitation pulse.[57] Ions are then irradiated with FELIX, which produces 5 μs long IR pulses with energies of 30-40 mJ inside the FTICR cell and a bandwidth of ≈0.5% of the central wavelength. IR spectra were recorded from 5 to 12.5 μm corresponding to a frequency range of 800 – 2000 cm^{-1}. A mass spectrum is generated using an excite/detect sequence as implemented in the FTICR control software adapted from that of Heeren and co-workers.[58] Three mass spectra were averaged at each wavelength point.

IRMPD occurs whenever the frequency of the photon is in resonance with the frequency of a molecular vibration; tens to hundreds of photons are absorbed leading to unimolecular dissociation.[32,34,59] From fragment and parent ion intensities observed in the mass spectrum, the fragmentation yield is derived as $\sum I$(fragments)/ $\sum I$(all ions). Plotting this yield as a function of the wavenumber of the radiation generates the infrared spectrum. The yield is corrected for variations in the laser pulse energy over the scanned wavelength range assuming a linear power dependence.

The b_2 fragment ions from PheGlnAla and AsnAlaAla are generated by nozzle–skimmer dissociation in the high-pressure region at the inlet of the mass spectrometer. Dissociation of AlaAsnAla is induced in the ICR cell by 0.1 s irradiation with a 35-W CO_2 laser (Universal Laser Systems, Scottsdale, AZ, USA, ULR-25), which was found to be more efficient than CID for this system. It was checked experimentally that fragment ions generated by either method exhibit the same IR spectrum.

7.2.2. Density functional theory calculations

Experimental spectra are compared to spectra computed for candidate structures at the density functional theory (DFT) level in order to identify the molecular structure of the ions generated. Optimized ion structures and their vibrational spectra were computed using the B3LYP functional and the 6-31++G(d,p) basis set, as implemented in Gaussian03, Revision C.02. Input structures were generated based on chemical intuition, where structures exhibiting resonance stabilization by conjugation, stabilization by hydrogen bonding, and avoiding syn steric interactions were particularly considered. A molecular dynamics search of the potential energy surfaces was not deemed necessary as the systems under study (particularly the b_2 fragment ions) are fairly small; moreover, the main goal of this study was to establish the isomeric structure and protonation site rather than the precise conformeric structure. Computed vibrational frequencies were scaled by 0.975 and convoluted with a 25 cm^{-1} full-width-at-half-maximum Gaussian line shape function to facilitate comparison with experimental spectra.

7.3. Results and discussion

7.3.1. Structure and protonation site of the PheGlnAla precursor peptide

Recent spectroscopic investigations of His, Arg, and Lys containing peptides have revealed their high propensity of forming diketopiperazine b_2-ions.[43-45] Protonation at the side chain nucleophile instead of at the N-terminal amine has been suggested to enhance the trans-to-cis isomerization of the peptide bond required for the formation of the diketopiperazine structure.[45] In other words, deviation from the formation of the classical oxazolone b_2-ion may thus be related to the protonation site in the precursor peptide. Since Gln and Asn containing peptides have been suggested to not form oxazolone b_2-ions [7,17,49-51], we first address the question of where the proton resides in the precursor PheGlnAla peptide.

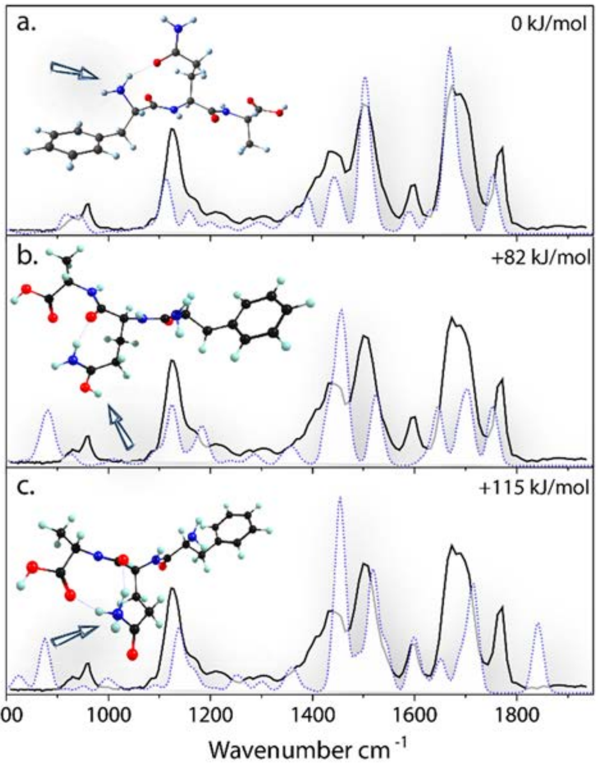


Figure 7.1. Comparison of the IRMPD spectrum of the protonated precursor peptide PheGlnAla (black trace) with DFT-calculated spectra (blue dashed lines) for structures protonated at the N-terminal nitrogen (panel a), Gln side-chain oxygen (panel b) and the Gln side-chain nitrogen (panel c). Calculated relative free energies are given for each structure. The arrow indicates the protonation site.

Although peptides without basic residues are generally assumed to be protonated on the N-terminus, protonation of backbone amide oxygen atoms has been suggested to be competitive.[60,61] Moreover, the Gln side-chain amine group may sequester the proton.

The main IRMPD fragment ions of protonated PheGlnAla are b_2 (m/z 276), $b_2 - NH_3$ (m/z 259), c_1 (m/z 165), z_2 (m/z 201) and a_1(m/z 120). The rather unusual observation of c_1 and z_2 fragments was also reported in CID studies and interpreted as evidence for non-oxazolone structures of the b-fragment.[50,51] Figure 7.1 compares the experimental IRMPD spectrum with computed spectra for PheGlnAla protonated at the N-terminus, at the glutamine side chain oxygen and at the glutamine side chain nitrogen (panels **a**, **b** and **c**, respectively). It is seen immediately that the match between experiment and theory is much better for protonation at the N-terminus. Moreover, the computed relative energies strongly favor protonation at the N-terminus.

We can thus safely assume that prior to activation, the peptide is protonated on the N-terminal amino group rather than on the Gln side chain, similar to what is observed for the majority of small protonated peptides and markedly different from peptides containing His, Arg, or Lys.

7.3.2. b_2-ion of protonated PheGlnAla

The b_2-ion of protonated PheGlnAla was produced by CID in the front end of the instrument and subsequently mass isolated in the FTICR-MS. Upon IRMPD, the ion undergoes dissociation forming charged products at m/z 120 (a_1) and m/z 165 (c_1), with the former being most abundant. This dissociation pathway is markedly different from that observed for typical oxazolone b_2-ions, which are known to generate mainly the a_2 ion via the loss of a carbon monoxide unit. It also differs from the most abundant IRMPD channels observed for diketopiperazine b_2 structures (i.e. cyclic dipeptides), which are mainly loss of CO, loss of CO+NH_3, and generation of the a_1 fragment.[37-39]

Figure 7.2 compares the experimental IRMPD spectrum of the PheGln b_2-ion at m/z 276 with calculated spectra for selected candidate structures. Ball-and-stick structures revealing the conformeric structures of these systems can be found in the Supplementary Material from Ref[1]. The diketopiperazine structure protonated on the carbonyl oxygen adjacent to the α-carbon of the Phe residue (panel **a**) is the global minimum structure. The structure with a 6-membered glutarimide ring resulting from a nucleophilic attack by the Gln side chain N-atom (panel **b**) is 4 kJ/mol higher in free energy, but still lower than the common oxazolone motif, protonated at the oxazolone nitrogen, at 50 kJ/mol (panel **c**). The alternative oxazolone structure, where the proton is at the N-terminus (panel **d**) is nearly iso-energetic. The imino-δ-valerolactone structure (panel **e**) resulting from nucleophilic attack by the Gln side chain O-atom, as suggested based on MS^3 data,[49,50] lies at more than 100 kJ/mol higher in energy. Protonation at the glutarimide or imino-valerolactone ring nitrogen atoms leads to structures that are substantially higher in free energy, such as the glutarimide protonated structure in panel **f**. Further protonation sites for each of these four main isomeric motifs were also investigated, but in all cases resulted in calculated vibrational spectra that deviate strongly from the experimental spectrum (not shown).

Inspection of the six calculated spectra in Figure 7.2 and comparison with the experimental spectrum leads quite readily to the conclusion that only the glutarimide structure in panel **b**, resulting from nucleophilic attack by the glutamine side chain nitrogen, can reasonably explain the experimental spectrum. An assignment of the most intense peaks in the

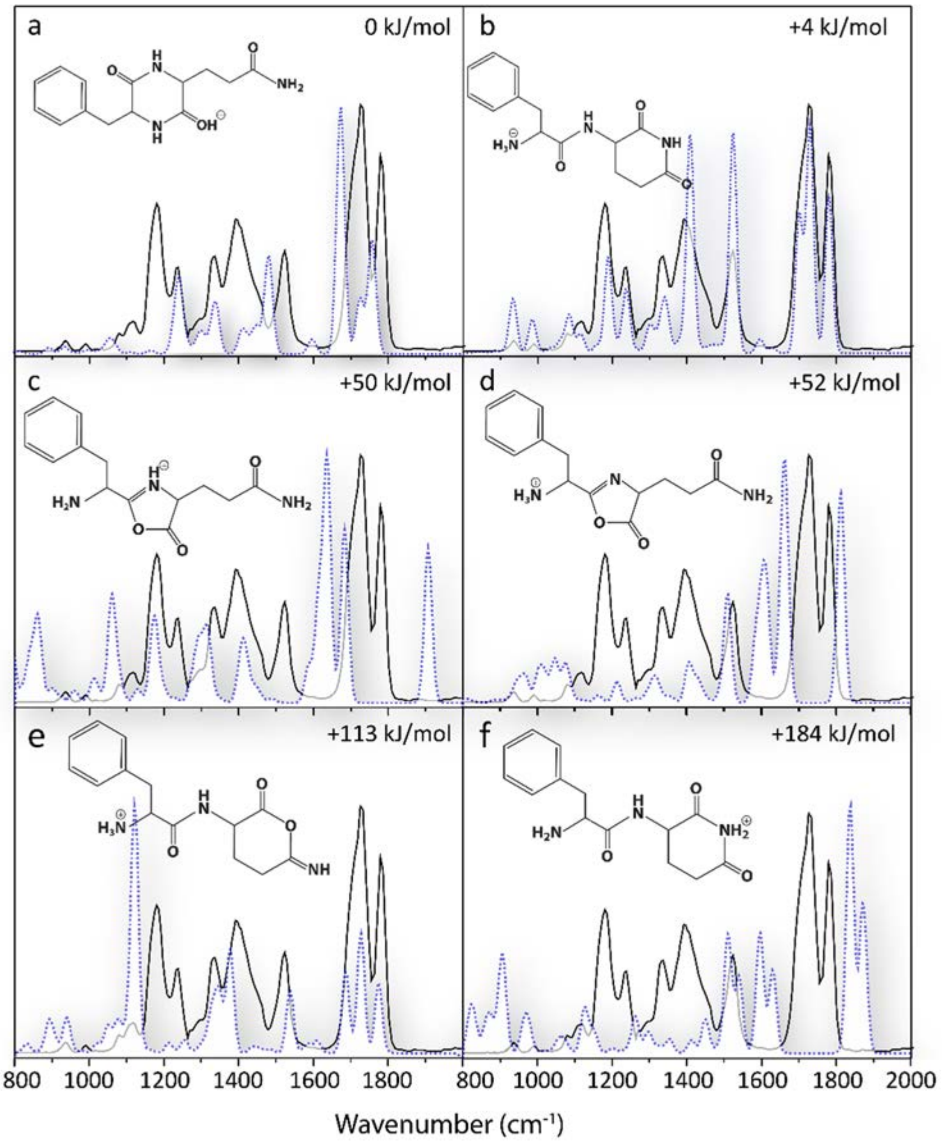


Figure 7.2. IRMPD spectrum of the b_2-ion of protonated PheGlnAla (black in all panels) compared to DFT-calculated spectra (blue dashed lines) for isomeric candidate structures: diketopiperazine protonated on carbonyl oxygen (a), glutarimide protonated at N-terminus (b), oxazolone protonated on oxazolone nitrogen (c), oxazolone protonated on N-terminus (d), 6-imino-δ-valerolactone protonated on N-terminus (e) and glutarimide protonated on glutarimide nitrogen (f). Calculated relative free energies are given for each of the structures. Three-dimensional ball-and-stick structures can be found in the Supplementary Material of Ref[1].

experimental IR spectrum is presented in Table 1. The carbonyl stretching modes found roughly between 1600 and 1800 cm^{-1} are, as in many cases, very characteristic and most diagnostic for the structural assignment. Three strong bands appear to be observed in this range, although two of them are not resolved. Of the computed spectra in Figure 7.2, only those of the glutarimide (**b**) and the imino-δ-valerolactone (**e**) structures reproduce the positions of the three bands. All other structures display strong bands in this region which are not observed experimentally, so that oxazolone as well as diketopiperazine structures can be confidently excluded. Based mainly on the comparison of experimental and theoretical spectra in the 1100 – 1300 cm^{-1} range, it is then concluded that the glutarimide structure matches the experimental spectrum substantially better than the isomeric imino-δ-valerolactone structure. However, small discrepancies between experimental and theoretical spectra are apparent in Figure 7.2b, particularly for the band at 1391 cm^{-1}. The band at 1391 cm^{-1} is assigned to the NH_3 umbrella vibration (computed at 1405 cm^{-1}), which is often found to be hard to reproduce accurately by computation, likely due the anharmonic shape of the potential well for these type of vibrations.[62-64]

While the glutarimide structure of Figure 7.2b clearly provides the best spectral match with the experiment, small deviations in the 900 – 1100 cm^{-1} range make a further experimental confirmation for this structural assignment desirable. First, we investigate the structure of the m/z 165 fragment ion, which has been assumed to be a c_1 ion.[49-51] Appearance of c-type ions is uncommon in CID mass spectra and has been interpreted as evidence for Gln side-chain induced rearrangement[49,50] (although it cannot distinguish between cyclic imide and isoimide structures; see Scheme 7.1).

Although c-type ions are rarely observed in CID of protonated peptides, they are well-known fragmentation products in electron capture and electron transfer dissociation (ECD/ETD)[6,65-68] and their structures have been extensively discussed.[8,9,69,70] In CID as well as in ECD/ETD, c-type ions result from cleavage of the N-C_α bond and are assumed to be linear N-terminal peptide fragments possessing either an amide or an enol-imine moiety at the C-terminus. Here, the IRMPD spectrum of the m/z 165 fragment ion of protonated PheGlnAla was obtained by recording the wavelength-dependent fragmentation into m/z 120, presumably the a_1 ion. Figure 7.3 compares the experimental IRMPD spectrum with computed spectra for the c_1 ion having either an amide (panel **a**) or an enol-imine (panel **b**) moiety at the C-terminus. Apart from the relative band intensities of a few of the minor bands, the experimental spectrum matches favorably with that calculated for the c_1-ion having a C-terminal amide, which is also substantially lower in energy than the isomeric enol-imine structure. The comparison in Figure 7.3a further suggests that the system is protonated on the N-terminus. The conclusion that the m/z 165 fragment ion is indeed the c_1 ion as previously suggested [17,49,50] thus lends support to the b_2-ion having either the glutarimide or the imino-δ-valerolactone structure.

Theory (cm^{-1})	Experimental (cm^{-1})	Vibrational Mode
1778	**1778**	glutarimide C=O stretch
1726	**1726**	glutarimide C=O stretch adjacent to α-C
1697	**1700 (shoulder)**	peptide bond C=O stretch
1521	**1523**	N-H bending / N-C stretch peptide bond
1405	**1391**	NH_3 umbrella motion
1335	**1331**	C-N-C sym stretch glutarimide ring
1232	**1235**	C-N-C asym stretch glutarimide ring
1181	**1180**	C-C stretch / CH_2 wag

Table 7.1. Experimental IR frequencies compared with calculated peak positions for the glutarimide structure of Figure 7.2b. Approximate normal mode descriptions are given in the last column.

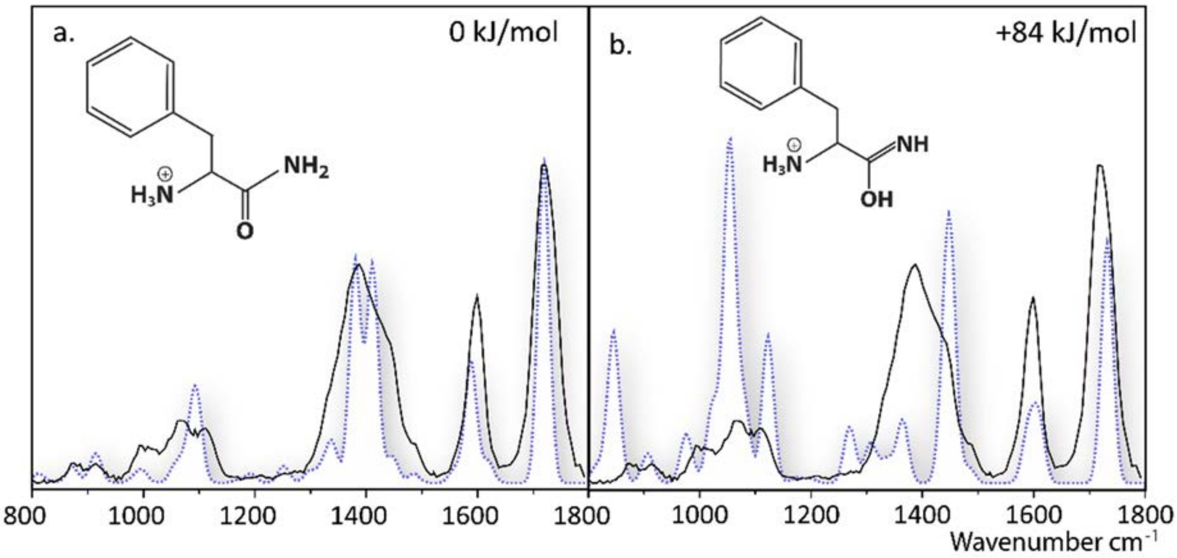


Figure 7.3. Comparison of the IRMPD spectrum of the m/z 165 fragment ion of protonated PheGlnAla (black trace) with DFT-calculated spectra for amide (panel **a**) and enol-imine (panel **b**) structures of the c_1 ion, both protonated at the N-terminal amino group (blue dashed lines). Relative free energy values are given. 3-dimensional structures are available in the Supplementary Material of Ref[1].

Secondly, to provide further support for our assignment as a glutarimide structure versus the imino-δ-valerolactone structure, an IRMPD spectrum was recorded for the conjugate acid of (S)-2-amino-N-((S)-2,6-dioxopiperidin-3-yl)-3-phenylpropanamide (see Figure 7.4), to serve as a reference for the assumed glutarimide b_2-ion structure. The compound was synthesized following the method of Fox et al.[55] and introduced into the FTICR-MS as a protonated ion by ESI. The experimental IRMPD spectrum is displayed in Figure 7.4, overlaid on the experimental spectrum of the b_2 fragment ion of protonated PheGlnAla. Within experimental accuracy, the two spectra are identical and moreover, the same dissociation channels are observed upon IRMPD. The excellent match in Figure 7.4 identifies the glutarimide structure as the sole contributor to the spectrum of the PheGlnAla b_2-ion, thus suggesting that the nucleophilic attack leading to the formation of the b_2-ion occurs through the Gln side chain nitrogen atom and not through its oxygen atom as suggested based on MS^3 data.[49,50]

Based on the experimental evidence, we can thus safely conclude that the PheGln b_2-ion has a glutarimide structure, which suggests that it is the side chain amide nitrogen atom that attacks the backbone carbonyl carbon. This is quite surprising as the side chain amide oxygen is considered to be a stronger nucleophile, since conjugation in the amide moiety places a partial negative charge on the O-atom and a partial positive charge on the N-atom. Furthermore, protonation at the O-atom of acetamide (CH_3-C(=O)-NH_2) is computed to be 54 kJ mol^{-1} more favorable than protonation at the N-atom.[71] In addition, the oxygen atoms in the amide linkages of peptides are known to be more basic than the nitrogen atoms.[60] Based on these considerations, it is perhaps not surprising that observation of the c_1-ion in MS^3 studies was interpreted as evidence for the formation of cyclic isoimides rather than cyclic imides by several authors.[7,49-51]

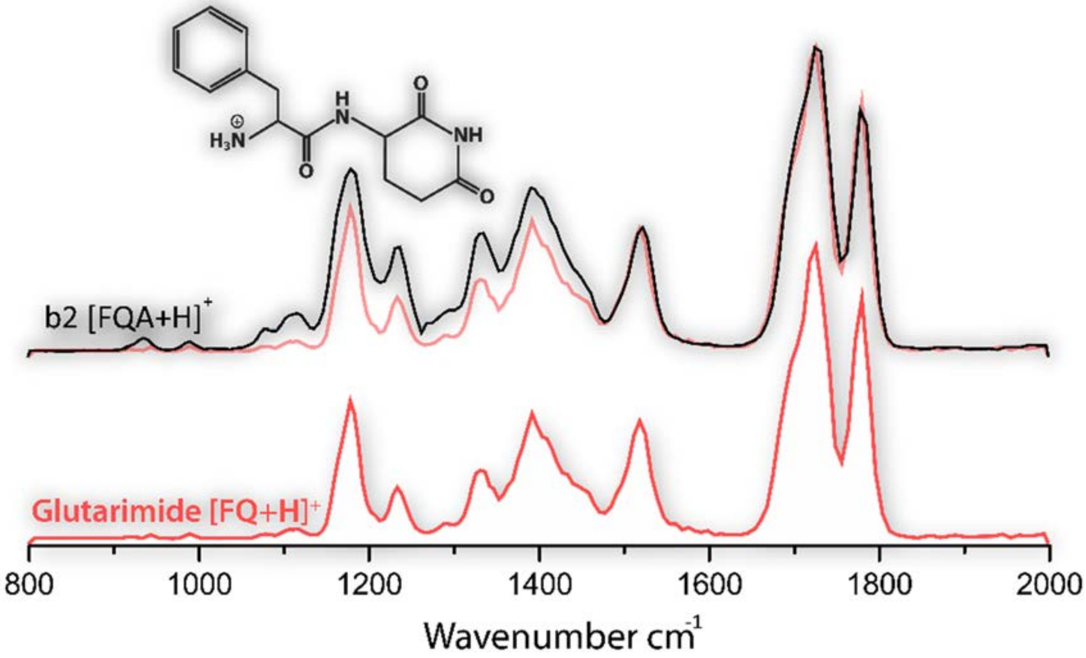


Figure 7.4. Comparison of the experimental IRMPD spectra of the protonated glutarimide containing reference compound (protonated (S)-2-amino-N-((S)-2,6-dioxopiperidin-3-yl)-3-phenylpropanamide, black) with that of the b_2 fragment ion from protonated PheGlnAla (red). All bands are seen to be reproduced exactly; the slight deviation in relative intensities is probably due to a slow decline of laser power during the scan of the b_2-ion, which is not accounted for in the power profile used for normalization.

To address the question why it is then the amide nitrogen that acts as the nucleophile and not the amide oxygen, we first consider the relative thermochemistry of the resulting product structures. Cyclization through the nitrogen atom leads to the glutarimide structure, which is about 80 kJ mol^{-1} (see Figure 7.2) more stable than the isomeric imino-δ-valerolactone structure resulting from cyclization through the oxygen atom. A reaction under thermodynamic rather than kinetic control could have explained the observed product,[72] if the diketopiperazine structure had not been lower in energy still. Note that the energy difference between the imide and isoimide structures is due to the isomerism of the

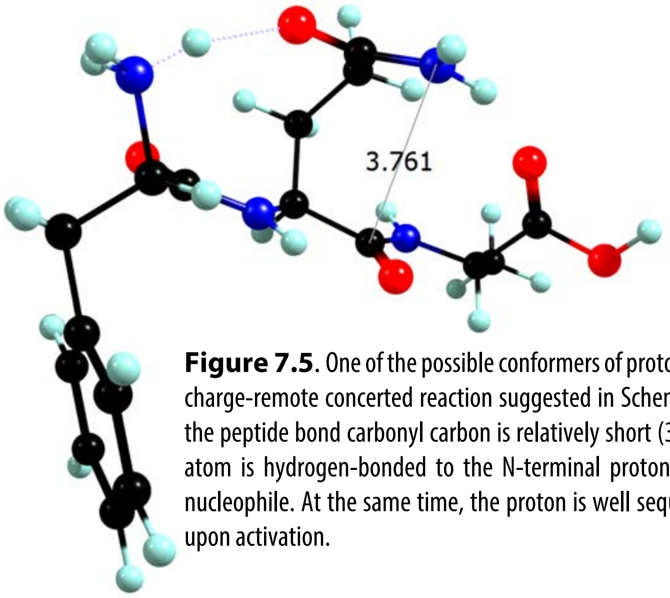
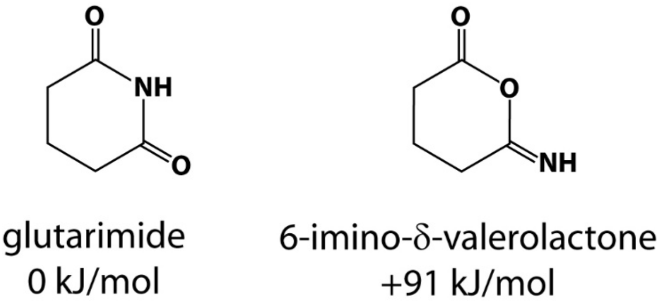


Figure 7.5. One of the possible conformers of protonated PheGlnAla that could serve as a starting point for the charge-remote concerted reaction suggested in Scheme 7.5. The distance from the glutamine nitrogen atom to the peptide bond carbonyl carbon is relatively short (3.761 Å). This structure also illustrates how the Gln oxygen atom is hydrogen-bonded to the N-terminal proton, thereby fixing this moiety so that it cannot act as the nucleophile. At the same time, the proton is well sequestered by two strong nucleophiles, reducing its mobility upon activation.

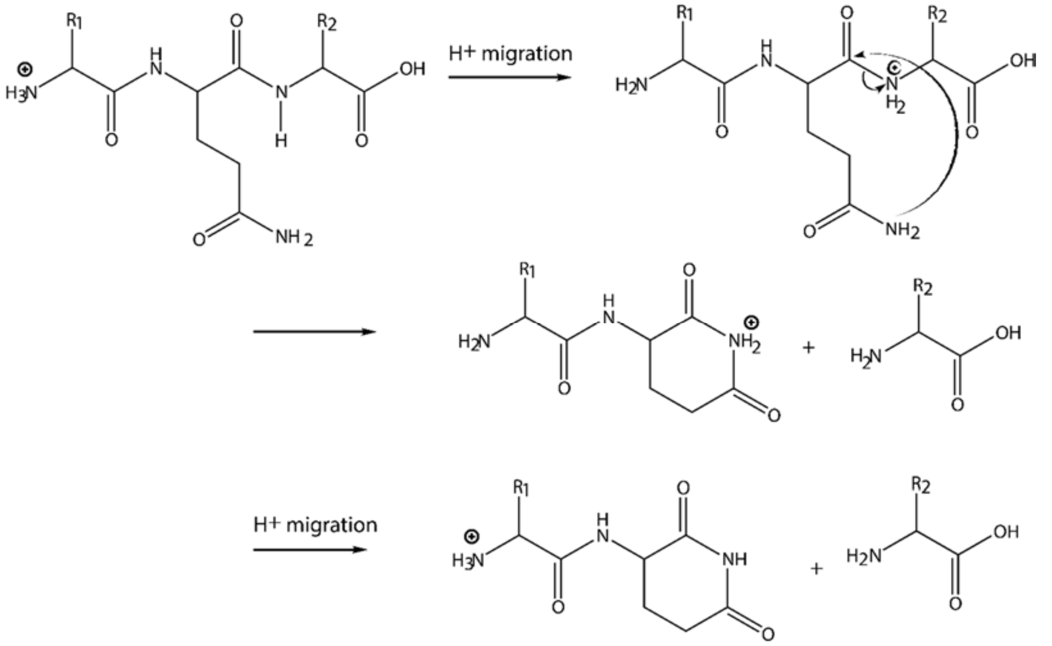


Scheme 7.3: Imide and isoimide isomers of $C_5H_7NO_2$ with their computed relative free energies.

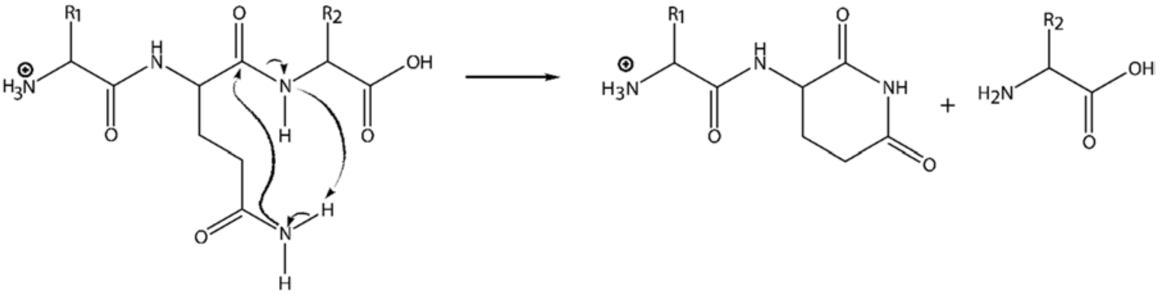
six-membered ring, and not to the conformational structure of the peptide chain; at the B3LYP/6-31++G** level of theory, the glutarimide isomer of $C_5H_7NO_2$ lies 91 kJ mol^{-1} lower in free energy than the imino-δ-valerolactone isomer (see Scheme 7.3). We note that O'Hair and coworkers[17] also investigated the relative thermochemistry of the two forms of the b$_2$-ion, although their peptides were N-acetylated. Therefore, only structures protonated at the (iso)imide ring were considered, so that their results are not entirely comparable to ours, where the proton resides on the N-terminal amino group of the b$_2$- fragment. For the b$_2$-fragments with a free N-terminus, protonation at the N-terminal amino group is computed to be substantially lower in energy (see Figure 7.2). Understanding the mechanistic details of the dissociation/rearrangement reaction requires extensive transition-state calculations, which are beyond the scope of this study. Speculatively, the reaction may not proceed via a mobile proton scenario (Scheme 7.4) but instead follow a charge-remote pathway, such as sketched in Scheme 7.5, similar to what has been suggested for glutamic acid (Glu) containing peptides[73] and similar to degradation pathways of Gln containing peptides in biological samples.[48] In the lowest-energy conformers of the precursor peptide (see Fig. 7.1 and Fig. 7.5), the protonated N-terminus is hydrogen-bonded to the side-chain amide O-atom of the Gln residue. This reduces the nucleophilicity of the side-chain O-atom and simultaneously sequesters the proton to some extent, possibly allowing for such charge-remote reactions to occur.

7.3.3. b$_2$-ion of protonated AlaAsnAla

Asparagine (Asn) differs from glutamine only in the length of the alkyl chain connecting the α-carbon and the amide moiety, which consists of one methylene unit for Asn versus two for Gln. O'Hair and coworkers suggested similar side-chain induced reactions could take place upon formation of b-type fragment ions for Asn containing peptides,[17] forming in this case five-membered ring species as shown in Scheme 7.2. However, Lee and Lee report that no c$_1$-ion is observed from peptides with Asn in the second position, suggesting that the Asn side chain does not interfere with amide bond cleavage.[50] In sharp contrast, Lehmann and coworkers report formation of c$_1$-ions from various peptides with Asn in the second position, though at lower abundance than for the Gln-containing peptides.[51] Relative thermodynamics of the side-chain rearranged species as compared to regular oxazolone species were suggested to be different for Gln and Asn.[50] Moreover, one could imagine the kinetics of the reaction to be different, for instance due to the shorter alkyl chain in Asn



Scheme 7.4. Possible dissociation/rearrangement reactions leading to glutarimide b$_2$ ion based on the mobile proton model. Migration of the added proton from the N-terminal amine group to the second amide linkage of the peptide is followed by nucleophilic attack of the glutamine nitrogen atom onto the peptide carbonyl carbon, forming the glutarmide protonated at the glutarimide nitrogen. The spectroscopically identified glutarimide structure is then formed by proton migration to the N-terminus.



Scheme 7.5. Alternative dissociation/rearrangement reaction leading to the glutarimide b$_2$ ion structure based on a charge-remote process. Without proton migration, a concerted process leads to the formation of a glutarimide ring. Figure 7.5 below shows a conformation of the precursor protonated PheGlnAla that could facilitate such a charge-remote dissociation reaction.

possibly restricting the accessibility of the amide heteroatoms to the adjacent carbonyl carbon. We therefore investigate here the structure of the b$_2$-ion from protonated AlaAsnAla to address the question whether the side chain of Asn becomes involved in the rearrangement reaction as observed for Gln or whether peptide dissociation proceeds via the more common oxazolone (or diketopiperazine) pathway.

In addition to the formation of oxazolone and diketopiperazine b$_2$-structures analogous to those depicted in Scheme 7.1, the Asn side-chain may become involved in the rearrangement reaction via one of the reactions outlined in Scheme 7.2, leading to two possible five-membered ring structures.[17] Nucleophilic attack by the side-chain amide nitrogen leads to the formation of a succinimide ring species, analogous to the glutarimide species observed for the Gln containing peptide. Alternatively, the b$_2$-ion may adopt a structure containing a 5-imino-γ-butyrolactone ring, if the nucleophile is the side chain amide oxygen atom. Depending on the protonation site, diketopiperazine and succinimide structures are iso-energetic within a few kJ mol^{-1} (see Figure 7.6). Oxazolone structures are typically around 30 kJ mol^{-1} higher in free energy, while structures containing a 5-imino-γ-butyrolactone ring are more than 80 kJ mol^{-1} higher, in accordance with values reported previously.[50] Note that this energetic ordering is significantly different from that for the analogous PheGln b_2-ion isomers in Figure 7.2.

Upon CO_2 laser induced multiple-photon dissociation, protonated AlaAsnAla was mainly observed to yield fragments at m/z 186 (b_2), m/z 258 (loss of ammonia), and m/z 257 (loss of water). At longer irradiation times, the c_1-ion becomes clearly observable, suggesting that the fragmentation pathways are (at least partly) analogous to those seen for the Gln containing peptide. However, direct formation of the c_1-ion, i.e. not traversing the b_2-ion, may also occur as suggested in Refs.[51,52]. An IRMPD-based MS^3 experiment on the b_2 fragment does not produce the a_2-ion as would have been expected for oxazolone or diketopiperazine b_2-ion structures, but instead produces mainly the a_1 fragment ion (m/z 44). While the c_1-ion is not as abundantly observed as for PheGln b_2-ion, formation of the a_1-ion is similar.

Theory (cm^{-1})	**Experimental (cm^{-1})**	**Vibrational Mode**
1820	**1836**	Succinimide C=O sym stretch
1768	**1764**	Succinimide C=O asym stretch
1695	**1725**	peptide bond C=O stretch
1528	**1533**	N-H bending / N-C stretch peptide bond
1398	**1432**	NH_3 umbrella motion
1340	**1340**	C-H bending modes
1300	**1315**	C-N-C sym. stretch succinimide ring
1242	**1250**	N-H & C-H wagging mode
1139	**1142**	C-N-C asym. stretch succinimide ring
1085	**1100**	N-C_α stretch
970	**966**	NH_3 & CH_3 twisting modes
890	**894**	C-C stretch / C-H twist

Table 2. Assignment of bands in the IR spectrum of the b_2 fragment of protonated AlaAsnAla based on the succinimide structure shown in Figure 7.6c.

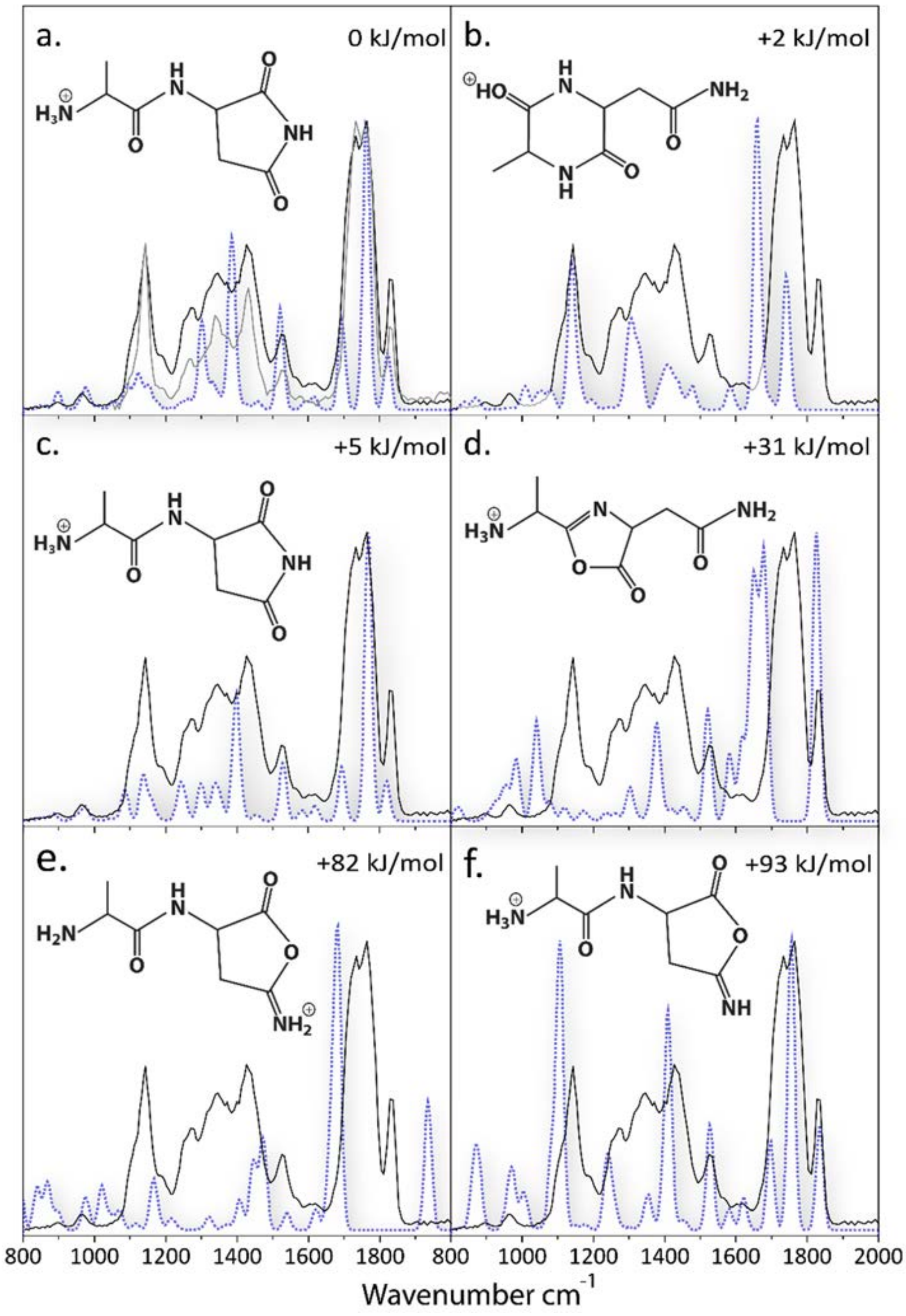


Figure 7.6. Comparison of the IRMPD spectrum for the b_2–ion of protonated AlaAsnAla (black trace in all panels) and DFT-calculated spectra (blue dashed line) for candidate structures: succinimide protonated at N-terminus (a,c), diketopiperazine protonated at carbonyl oxygen (b), oxazolone protonated on N-terminus (d), 5-imino-γ-butyrolactone protonated on imino nitrogen (e) and 5-imino-γ-butyrolactone protonated on N-terminal nitrogen (f).The succinimide structures in panels a and c are isomerically identical but conformationally different, as shown by the ball-and-stick models. Calculated relative free energy values are indicated for each of the structures. 3-dimensional structures are available in the Supporting Information of Ref[1].

The experimental IRMPD spectrum of the AlaAsn b_2 fragment ion is compared with theoretical spectra for six representative structural isomers in Figure 7.6. Further structures differing in protonation site or conformation not shown in Figure 7.6 have higher relative energies and spectra poorly matching experiment. The figure displays calculated spectra for the succinimide ring structure protonated at the N-terminus (two different conformers in panels **a** and **c**), the diketopiperazine ring protonated at one of the carbonyls (panel **b**), the oxazolone structure protonated at the N-terminus (panel **d**), the structure having a 5-imino-γ-butyrolactone "isoimide" ring protonated at the imino nitrogen (panel **e**) and the same butyrolactone structure protonated on the N-terminal amino group (panel **f**). Panel **a** shows in addition a scan taken at reduced laser power. 3-dimensional structures revealing the conformations are shown in the Supporting Information Ref[1].

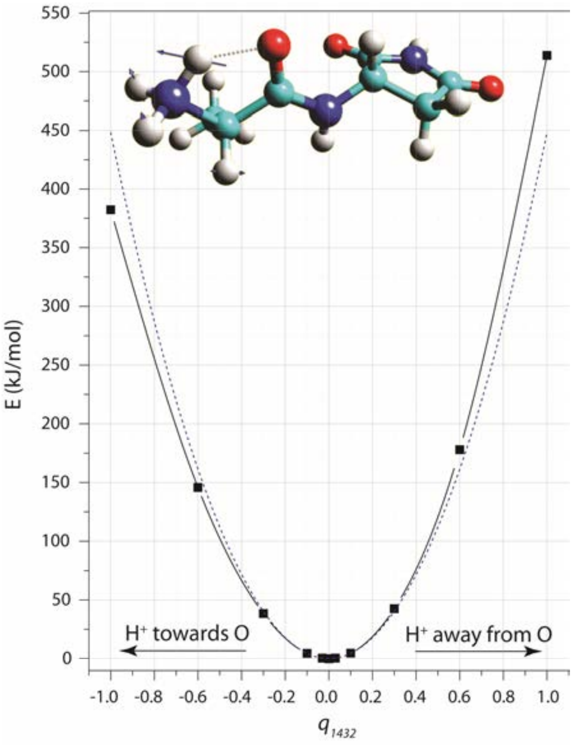


Figure 7.6. Potential energy scan (solid black) following the normal mode corresponding to the 1432 cm^{-1} vibration in the succinimide structure of the b_2-ion of protonated AlaAsnAla. The dotted line represents a perfect harmonic curve. The normal mode is seen to involve mainly NH bending of the strongly hydrogen-bonded H-atom.

Again, the carbonyl stretching region between 1600 and 1900 cm^{-1} is very diagnostic. The trident peak shape observed in the experimental spectrum is only reproduced by the succinimide and butyrolactone structures. Diketopiperazine (panel **b**) and oxazolone (panel **d**) structures can be confidently excluded, since they utterly fail to match with the experiment in this diagnostic region of the spectrum. Protonation at the succinimide or imino-butyrolactone nitrogen atoms (e.g. panel **e**) can also be excluded based on these grounds, which is in line with the computed relative energies: protonation at the N-terminus is substantially favored over protonation at the imide or isoimide ring. Calculated spectra for both the succinimide (panel **a** and **c**) and 5-imino-γ-butyrolactone (panel **f**) structures, protonated at the N-terminus, show reasonable overlap in the 1600-1900 cm^{-1} spectral range. However, peak positions in the 800-1200 cm^{-1} spectral range match much better with the succinimide structure than with the imino-butyrolactone structure; combined with the much higher relative free energy and analogous to the PheGln b_2-ion structure, we therefore discard the imino-butyrolactone isomer as possible structure for the b_2-ion of AlaAsnAla.

The computed spectra for the two conformers (panels **a**, **c**) of the succinimide-based structure match the experimental spectrum reasonably well except for the region between 1200 and 1500 cm^{-1}. A comparison between experimentally observed frequencies and those calculated for the conformer in panel **c** is given in Table 2. It is of interest to note here that the same spectral region was also poorly reproduced by the calculations for the glutarimide structure for the PheGln b_2 fragment ion (vide supra). The largest deviation between experiment and theory is found for the NH_3 umbrella mode calculated at 1398 cm^{-1} and tentatively assigned to the strong band observed at 1432 cm^{-1}. Following the potential energy along this normal coordinate yields the curve plotted in Figure 7.6, exhibiting the anharmonic nature of the potential, which is likely also due to the strong hydrogen bond to the peptide carbonyl O-atom. Note that the C=O stretching mode of this peptide bond is also miscalculated substantially at 1695 cm^{-1} versus an observed frequency of 1725 cm^{-1}. Despite the mismatch between experimental and DFT-calculated spectra in the 1200-1500 cm^{-1} range, we conclude that the succinimide isomer is the major contributor to spectrum observed for the AlaAsn b_2 fragment ion. This suggests that the Asn side-chain nitrogen atom acts as the nucleophile on the reaction pathway towards this fragment, analogous to what is observed for the Gln residue. This observation appears somewhat in contrast with that of Ref.[50], where peptides with Gln and Asn in the second position were suggested to behave differently in the formation of the b_2-ion.

7.3.4. b_2-ion of protonated AsnAlaAla

It has been suggested that b_n-ions of longer protonated peptides with Gln or Asn in the n-th position may give rise to structures similar to those observed for the b_2-ions.[50] While such longer peptide fragments are beyond the scope of this investigation, it is of interest to investigate the structure of a b_2-ion from a peptide with Gln or Asn in the first position. Will the side chain amide nitrogen continue to be the nucleophile leading to cyclic imide structures, or will the b_2-ions adopt one of the more common motifs? MS^3 experiments report b_2-ions from peptides with Asn or Gln in the first position to yield predominantly the a_2-ion,[51] presumably formed by loss of carbon monoxide from a classical oxazolone b_2-ion. Here we investigate the influence of an amide containing residue in the first position by recording the IRMPD spectrum for the b_2-ion of protonated AsnAlaAla.

Upon CID, this peptide (m/z 275) was found to undergo dissociation mainly into m/z 257 (loss of water), m/z 258 (loss of ammonia), b_2 (m/z 186) and a_2 (m/z 158). No c- or z-type ions were observed in accordance with previous CID experiments.[51] The b_2-ion was mass isolated and underwent fragmentation into m/z 44 (NH_3-$CH=CH_2^+$) and the a_2–ion at m/z 169. This breakdown pathway resembles that observed for the oxazolone AlaAla, AlaGly and GlyGly b_2 species.

The IRMPD spectrum of the b_2-ion of protonated AsnAlaAla indeed confirms an oxazolone structure as shown in Figure 7.7. Along with the experimental spectrum, calculated spectra are shown for oxazolone structures protonated either at the N-terminal nitrogen or at the oxazolone nitrogen, which are computed to be close in energy (5 kJ mol^{-1}). The diagnostic C=O stretching band of the oxazolone carbonyl at frequencies near 1900 cm^{-1} is

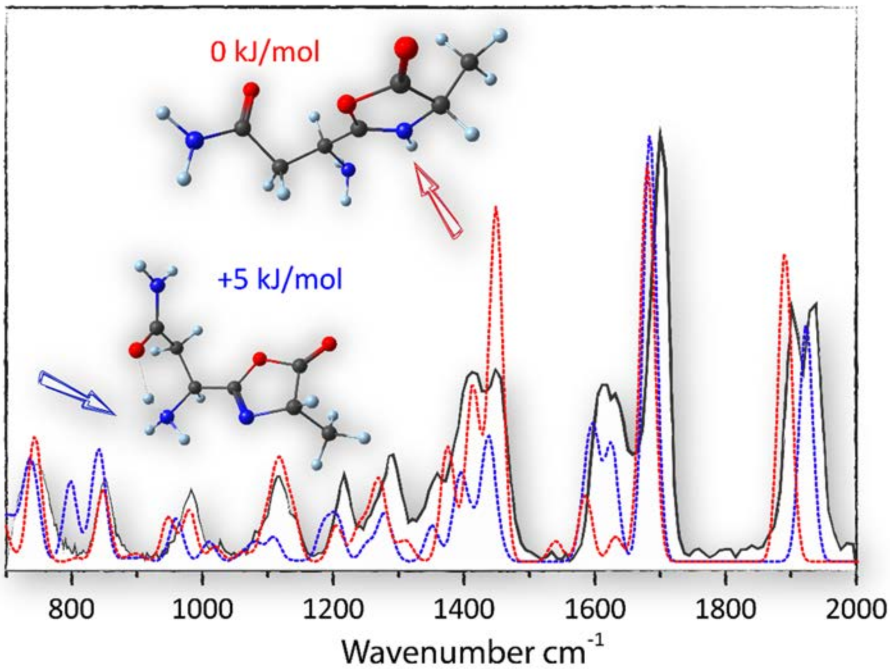


Figure 7.7. Comparison of the IRMPD spectrum for the b_2-ion of protonated AsnAlaAla (black trace) with DFT-calculated spectra for oxazolone structures protonated at the N-terminus (blue dashed line) and at the oxazolone nitrogen (red dashed line). The arrow indicates the protonation site.

prominently present and is seen to be split into two components. As seen from the comparison with the theoretical spectra, the splitting of the oxazolone band suggests the coexistence of both oxazolone structures. All bands in the spectrum are seen to be reproduced by the bands calculated for the two oxazolone structures, strongly suggesting that any contribution from alternative structures can be ruled out for this b_2-ion. Nucleophilic attack from either the amide side chain oxygen or nitrogen atom would give rise to cyclic 8-membered imide or isoimide structures. Their computed spectra indeed do not match with the experimental spectrum, as shown in the Supporting Information of Ref[1].

Observation of the N-terminally protonated oxazolone structure is somewhat in contrast with the AlaAla[38], AlaGly[37] and GlyGly[22] oxazolone b_2-ion structures, which were spectroscopically found to be protonated exclusively on the oxazolone nitrogen atom. On the other hand, coexistence of both protonation structures was also observed for the b_2-ion from protonated Leu-Enkephaline,[23,35,36] which was explained by additional stabilization of the higher-energy N-terminally protonated isomer by interactions with the Tyr π-cloud. For the AsnAla b_2-ion studied here, additional stabilization of the N-terminally protonated structure is provided by hydrogen bonding to the amide oxygen of the Asn residue.

7.4. Conclusions

Infrared ion spectroscopy was applied in combination with electronic structure calculations to resolve the molecular structure of b_2 peptide fragment ions containing a Gln or Asn residue. Previous MS^n based studies on peptides with Gln or Asn as the second residue had suggested that the amide side chain becomes involved in the dissociation reaction and that the b_2-ion forms a cyclic isoimide structure via nucleophilic attack by the side chain amide oxygen atom.[49-51] Our spectroscopic results presented herein show however that the ring formation occurs through the side chain amide nitrogen instead, forming b_2 fragments with glutarimide and succinimide structures for Gln and Asn, respectively. Although the match between experimental and computed harmonic IR spectra shows some deviations in the 1200-1500 cm^{-1} range, an experimental IR spectrum of a synthesized glutarimide reference compound agrees perfectly with that of the PheGln b_2-ion. The cyclic imide structures resulting from amide nitrogen nucleophilic attack are substantially lower in energy than the isomeric isoimide structures. Further computational investigation of the potential energy surface of the reaction are required to resolve the mechanistic details of the reaction, but is well beyond the scope of the present study. The structures determined here are expected to provide important guides in finding the correct trajectories.

As compared to the previously suggested isoimide structures, formation of imide b_2-ion structures is evenly consistent with available MS^3 data, in particular with the formation of a c_1-ion upon dissociation of the b_2-ion. In addition, IRMPD spectroscopy was applied here to confirm that the dissociation product indeed has the structure of a c_1-ion, with a C-terminal amide moiety and protonation at the N-terminus.

The protonation site of the PheGlnAla precursor peptide is determined to be the N-terminal amino group rather than the side chain amide moiety. The fact that no diketopiperazine structure is found here for the b_2-ion lends further support to the dissociation mechanism recently proposed for peptides with His, Arg and Lys residues.[43-45]

IR spectroscopy of the b_2-ion of protonated AsnAlaAla confirms that side-chain involvement in the dissociation reaction can be clearly ruled out in this case and that this Asn-containing fragment possesses purely an oxazolone structure. This conclusion was also inferred from MS^3 data,[51] although no distinction between oxazolone and diketopiperazine structures could be made. Detailed inspection of the IR spectrum recorded here reveals that structures protonated both at the oxazolone nitrogen as well as at the N-terminus coexist.

Acknowledgments

We gratefully thank Dr. Jan van Maarseveen and Stanimir Popovic from the University of Amsterdam for synthesizing the glutarimide model compound. We would like to thank the FELIX staff, in particular from Drs. A.F.G. van der Meer, B. Redlich and G. Berden, for their expert technical support. JO acknowledges the Stichting Physica for support. Financial support for this project has been provided by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) in the form of VICI grant no. 724.011.002. We thank the SARA Supercomputer Center for providing the computational resources. This work is part of the research program of FOM, which is financially supported by NWO.

REFERENCES

(1) Grzetic, J.; Oomens, J. *Spectroscopic identification of cyclic imide b_2-ions from peptides containing Gln and Asn residues* J. Am. Soc. Mass Spectrom. **2013**, *24*, 1228.

(2) Aebersold, R.; Goodlett, D. R. *Mass spectrometry in proteomics* Chem. Rev. **2001**, *101*, 269.

(3) Steen, H.; Mann, M. *The ABC's (and XYZ's) of peptide sequencing* Nat. Rev. Mol. Cell Biol. **2004**, *5*, 699.

(4) Burlet, O.; Orkiszewski, R. S.; Ballard, K. D.; Gaskell, S. *Charge promotion of low-energy fragmentations of peptide ions* Rapid Commun. Mass Spectrom. **1992**, *6*, 658.

(5) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *Mobile and localized protons: A framework for understanding peptide dissociation.* J. Mass Spectrom. **2000**, *35*, 1399.

(6) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *Electron capture dissociation of multiply charged protein cations. A nonergodic process* J. Am. Chem. Soc. **1998**, *120*, 3265.

(7) Paizs, B.; Suhai, S. *Fragmentation pathways of protonated peptides* Mass Spectrom. Rev. **2004**, *24*, 508.

(8) Turecek, F. *N-C$_\alpha$ bond dissociation energies and kinetics in amide and peptide radicals. Is the dissociation a non-ergodic process?* J. Am. Chem. Soc. **2003**, *125*, 5954.

(9) Syrstad, E. A.; Turecek, F. *Toward a general mechanism of electron capture dissociation* J. Am. Soc. Mass Spectrom. **2004**, *16*, 208.

(10) Barlow, C. K.; O'Hair, R. A. J. *Gas-Phase Peptide Fragmentation: How Understanding the Fundamentals Provides a Springboard to Develping New Chemistry and Novel Proteomics Tools* J. Mass Spectrom. **2008**, *43*, 1301.

(11) Harrison, A. G. *To b or not to b: the ongoing saga of peptide b ions* Mass Spectrom. Rev. **2009**, *28*, 640.

(12) Bythell, B. J.; Maitre, P.; Paizs, B. *Cyclization and Rearrangement Reactions of a_n Fragment Ions of Protonated Peptides* J. Am. Chem. Soc. **2010**, *132*, 14766.

(13) Polce, M. J.; Ren, D.; Wesdemiotis, C. *Dissociation of the Peptide Bond in Protonated Peptides* J. Mass. Spectrom. **2000**, *35*, 1391.

(14) Boyd, R.; Somogyi, A. *The mobile proton hypothesis in fragmentation of protonated peptides: a perspective* J. Am. Soc. Mass Spectrom. **2010**, *21*, 1275.

(15) Yalcin, T.; Khouw, C.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. *Why are b ions stable species in peptide spectra?* J. Am. Soc. Mass Spectrom. **1995**, *6*, 1164.

(16) Paizs, B.; Lendvay, G.; Vekey, K.; Suhai, S. *Formation of b_2^+ ions from protonated peptides: an ab initio study* Rapid Commun. Mass Spectrom. **1999**, *13*, 525.

(17) Farrugia, J. M.; OHair, R. A. J.; Reid, G. E. *Do all b_2 ions have oxazolone structures? Multistage mass spectrometry and ab initio studies on protonated N-acyl amino acid methyl ester model systems* Int. J. Mass Spectrom. **2001**, *210-211*, 71.

(18) Bleiholder, C.; Osburn, S.; Williams, T. D.; Suhai, S.; Van Stipdonk, M.; Harrison, A. G.; Paizs, B. *Sequence-Scrambling Fragmentation Pathways of Protonated Peptides* J. Am. Chem. Soc. **2008**, *130*, 17774.

(19) Harrison, A. G. *Peptide Sequence Scrambling Through Cyclization of b_5 Ions* J. Am. Soc. Mass. Spectrom. **2008**, *19*, 1776.

(20) Molesworth, S.; Osburn, S.; Van Stipdonk, M. *Influence of Size on Apparent Scrambling of Sequence During CID of b-Type Ions* J. Am. Soc. Mass. Spectrom. **2009**, *20*, 2174.

(21) Molesworth, S.; Osburn, S.; Van Stipdonk, M. *Influence of Amino Acid Side Chains on Apparent Selective Opening of Cyclic b(5) Ions* J. Am. Soc. Mass. Spectrom. **2010**, *21*, 1028.

(22) Chen, X.; Yu, L.; Steill, J. D.; Oomens, J.; Polfer, N. C. *Effect of peptide fragment size on the propensity of cyclization in collision-induced dissociation: oligoglycine b_2–b_8* J. Am. Chem. Soc. **2009**, *131*, 18272.

(23) Chen, X.; Steill, J. D.; Oomens, J.; Polfer, N. C. *Oxazolone versus macrocycle structures for Leu-Enkephalin b_2-b_4: Insights from infrared multiple-photon dissociation spectroscopy and gas-phase hydrogen/deuterium exchange* J. Am. Soc. Mass Spectrom. **2010**, *21*, 1313.

(24) Li, X.; Huang, Y.; O'Connor, P. B.; Lin, C. *Structural Heterogeneity of Doubly-Charged Peptide b-Ions* J. Am. Soc. Mass Spectrom. **2011**, *22*, 245.

(25) Erlekam, U.; Bythell, B. J.; Scuderi, D.; Van Stipdonk, M.; Paizs, B.; Maitre, P. *Infrared spectroscopy of fragments of protonated peptides: Direct evidence for macrocyclic structures of b_5 ions* J. Am. Chem. Soc. **2009**, *131*, 11503.

(26) Tirado, M.; Rutters, J.; Chen, X.; Yeung, A.; van Maarseveen, J.; Eyler, J. R.; Berden, G.; Oomens, J.; Polfer, N. C. *Disfavoring Macrocycle b Fragments by Constraining Torsional Freedom: The "Twisted" Case of QWFGLM b_6* J. Am. Soc. Mass Spectrom. **2012**, *23*, 475.

(27) Chen, X.; Tirado, M.; Steill, J. D.; Oomens, J.; Polfer, N. C. *Cyclic peptide as reference system for b ion structural analysis in the gas phase* J. Mass Spectrom. **2011**, *46*, 1011.

(28) Savitski, M. M.; Falth, M.; Fung, Y. M. E.; Adams, C. M.; Zubarev, R. A. *Bifurcating Fragmentation Behavior of Gas-Phase Tryptic Peptide Dications in Collisional Activation* J. Am. Soc. Mass Spectrom. **2008**, *19*, 1755.

(29) Paizs, B.; Suhai, S. *Combined Quantum Chemical and RRKM Modeling of the Main Fragmentation Pathways of Protonated GGG. I. Cis-Trans Isomerization Around Protonated Amide Bonds* Rapid Commun. Mass Spectrom. **2001**, *15*, 2307.

(30) Balta, B.; Aviyente, V.; Lifshitz, C. *Elimination of Water from the Carboxyl Group of $GlyGlyH^+$* J. Am. Soc. Mass Spectrom. **2003**, *14*, 1192.

(31) Armentrout, P. B.; Heaton, A. L. *Thermodynamics and mechanisms of Protonated Diglycine Decomposition: A Computational Stucy* J. Am. Soc. Mass Spectrom. **2012**, *23*, 621.

(32) Oomens, J.; Sartakov, B. G.; Meijer, G.; von Helden, G. *Gas-phase infrared multiple photon dissociation spectroscopy of mass-selected molecular ions* Int. J. Mass Spectrom. **2006**, *254*, 1.

(33) Eyler, J. R. *Infrared multiple photon dissociation spectroscopy of ions in Penning traps* Mass Spectrom. Rev. **2009**, *28*, 448.

(34) Polfer, N. C. *Infrared multiple photon dissociation spectroscopy of trapped ions* Chem. Soc. Rev. **2011**, *40*, 2211.

(35) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *Spectroscopic and theoretical evidence for oxazolone ring formation in collision-induced dissociation of peptides* J. Am. Chem. Soc. **2005**, *127*, 17154.

(36) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *Infrared spectroscopy and theoretical studies on gas-phase protonated Leu-enkephalin and its fragments: Direct experimental evidence for the mobile proton* J. Am. Chem. Soc. **2007**, *129*, 5887.

(37) Yoon, S. H.; Chamot-Rooke, J.; Perkins, B. R.; Hilderbrand, A. E.; Poutsma, J. C.; Wysocki, V. H. *IRMPD spectroscopy shows that AGG forms an oxazolone b_2^+ ion* J. Am. Chem. Soc. **2008**, *130*, 17644.

(38) Oomens, J.; Young, S.; Molesworth, S.; van Stipdonk, M. *Spectroscopic evidence for an oxazolone structure of the b_2 fragment ion from protonated tri-alanine* J. Am. Soc. Mass Spectrom. **2009**, *20*, 334.

(39) Wang, D.; Gulyuz, K.; Stedwell, C. N.; Polfer, N. C. *Diagnostic NH and OH Vibrations for Oxazolone and Diketopiperazine Structures: b2 from Protonated Triglycine* J. Am. Soc. Mass Spectrom. **2011**, *22*, 1197.

(40) Bythell, B. J.; Erlekam, U.; Paizs, B.; Maitre, P. *Infrared spectroscopy of fragments from doubly protonated tryptic peptides* Chem. Phys. Chem. **2009**, *10*, 883.

(41) Sinha, R. K.; Erlekam, U.; Bythell, B. J.; Paizs, B.; Maitre, P. *Diagnosing the Protonation Site of b_2 Peptide Fragment Ions using IRMPD in the X-H (X = O, N, and C) Stretching Region* J. Am. Soc. Mass Spectrom. **2011**, *22*, 1645.

(42) Morrison, L.; Somogyi, A.; Wysocki, V. H. *The Influence Glutamic Acid in Protonated b_3 --> b_2 Formation from VGEIG and Related Analogs* Int. J. Mass Spectrom. **2012**, *325-327*, 139.

(43) Perkins, B. R.; Chamot-Rooke, J.; Yoon, S. H.; Gucinski, A. C.; Somogyi, A.; Wysocki, V. H. *Evidence of diketopiperazine and oxazolone structures for HA b_2^+ ion* J. Am. Chem. Soc. **2009**, *131*, 17528.

(44) Zou, S.; Oomens, J.; Polfer, N. C. *Competition between diketopiperazine and oxazolone formation in water loss products from protonated ArgGly and GlyArg* Int. J. Mass Spectrom. **2012**, *316-318*, 12.

(45) Gucinski, A. C.; Chamot-Rooke, J.; Nicol, E.; Somogyi, A.; Wysocki, V. H. *Structural Influences on Preferential Oxazolone versus Diketopiperazine b_2^+ Ion Formation for Histidine Analogue-Containing Peptides* J. Phys. Chem. A **2012**, *116*, 4296.

(46) O'Hair, R. A. J. *The role of nucleophile–electrophile interactions in the unimolecular and bimolecular gas-phase ion chemistry of peptides and related systems* J. Mass. Spectrom. **2000**, *35*, 1377.

(47) In Ref. 16, the text and the values in Table 3 are ambiguous about the energetic ordering of the structures resulting from side-chain nitrogen and side-chain oxygen attack, referred to as Gln G1 and Gln G2, respectively.

(48) Jonsson, A. P.; Bergman, T.; Jörnvall, H.; Griffiths, W. J.; Bratt, P.; Strömberg, N. *Gln-Gly Cleavage: Correlation Between Collision-Induced Dissociation and Biological Degradation* J. Am. Soc. Mass Spectrom. **2001**, *12*, 337.

(49) Harrison, A. G. *Fragmentation reactions of protonated peptides containing glutamine or glutamic acid* J. Mass Spectrom. **2003**, *38*, 174.

(50) Lee, Y. J.; Lee, Y. M. *Formation of c1 fragment ions in collision-induced dissociation of glutamine-containing peptide ions: a tip for de novo sequencing* Rapid Comm. Mass Spectrom. **2004**, *18*, 2069.

(51) Winter, D.; Seidler, J.; Hahn, B.; Lehmann, W. *Structural and mechanistic information on c_1 ion formation in collision-induced fragmentation of peptides* J. Am. Soc. Mass Spectrom. **2010**, *21*, 1814.

(52) Winter, D.; Lehmann, W. D. *Sequencing of the Thirteen Structurally Isomeric Quartets of N-terminal Dipeptide Motifs in Peptides by Collision-Induced Dissociation* Proteomics **2009**, *9*, 2076.

(53) Oepts, D.; van der Meer, A. F. G.; van Amersfoort, P. W. *The infrared free-electron laser facility FELIX* Infrared Phys. Technol. **1995**, *36*, 297.

(54) Valle, J. J.; Eyler, J. R.; Oomens, J.; Moore, D. T.; van der Meer, A. F. G.; von Helden, G.; Meijer, G.; Hendrickson, C. L.; Marshall, A. G.; Blakney, G. T. *Free electron laser-Fourier transform ion cyclotron resonance mass spectrometry facility for obtaining infrared multiphoton dissociation spectra of gaseous ions* Rev. Sci. Instrum. **2005**, *76*, 023103.

(55) Fox, D. J.; Reckless, J.; Warren, S. G.; Grainger, D. J. *Design, Synthesis, and Preliminary Pharmacological Evaluation of N-Acyl-3-aminoglutarimides as Broad-Spectrum Chemokine Inhibitors in Vitro and Anti-inflammatory Agents in Vivo* J. Med. Chem. **2002**, *45*, 360.

(56) Polfer, N. C.; Oomens, J. *Reaction products in mass spectrometry elucidated with infrared spectroscopy* Phys. Chem. Chem. Phys. **2007**, *9*, 3804.

(57) Marshall, A. G.; Wang, T. C. L.; Ricca, T. L. *Tailored excitation for Fourier transform ion cyclotron mass spectrometry* J. Am. Chem. Soc. **1985**, *107*, 7893.

(58) Mize, T. H.; Taban, I.; Duursma, M.; Seynen, M.; Konijnenburg, M.; Vijftigschild, A.; Doornik, C. V.; Rooij, G. V.; Heeren, R. M. A. *A modular data and control system to improve sensitivity, selectivity, speed of analysis, ease of use, and transient duration in an external source FTICR-MS* Int. J. Mass Spectrom. **2004**, *235*, 243.

(59) Polfer, N. C.; Oomens, J. *Vibrational spectroscopy of bare and solvated ionic complexes of biological relevance* Mass Spectrom. Rev. **2009**, *28*, 468.

(60) Rodriguez, C. F.; Cunje, A.; Shoeib, T.; Chu, I. K.; Hopkinson, A. C.; Siu, K. W. M. *Proton Migration and Tautomerism in Protonated Triglycine* J. Am. Chem. Soc. **2001**, *123*, 3006.

(61) Wu, R.; McMahon, T. B. *Infrared Multiple Photon Dissociation Spectroscopy as Structural Confirmation for GlyGlyGlyH$^+$ and AlaAlaAlaH$^+$ in the Gas Phase. Evidence for Amide Oxygen as the Protonation Site* J. Am. Chem. Soc. **2007**, *129*, 11312.

(62) Sinclair, W. E.; Pratt, D. W. J. Chem. Phys. **1996**, *105*, 7942.

(63) Piest, H.; von Helden, G.; Meijer, G. *Infrared spectroscopy of jet-cooled neutral and ionized aniline--Ar* J. Chem. Phys. **1999**, *110*, 2010.

(64) Luis, J. M.; Reis, H.; Papadopoulos, M.; Kirtman, B. *Treatment of nonlinear optical properties due to large amplitude anharmonic vibrational motions: Umbrella motion in NH_3* J. Chem. Phys. **2009**, *131*, 034116.

(65) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations* Anal. Chem **2000**, *72*, 563.

(66) Cooper, H. J.; Hakansson, K.; Marshall, A. G. *The role of electron capture dissociation in biomolecular analysis* Mass Spectrom. Rev. **2005**, *24*, 201.

(67) Zubarev, R. A. *Reactions of polypeptide ions with electrons in the gas phase* Mass Spectrom. Rev. **2003**, *22*, 57.

(68) Coon, J. J.; Shabanowitz, J.; Hunt, D. F.; Syka, J. E. P. *Electron transfer dissociation of peptide anions* J. Am. Soc. Mass Spectrom. **2005**, *16*, 880.

(69) Frison, G.; van der Rest, G.; Turecek, F.; Besson, T.; Lemaire, J.; Maitre, P.; Chamot-Rooke, J. *Structure of electron-capture dissociation fragments from charge-tagged peptides probed by tunable infrared multiple photon dissociation* J. Am. Chem. Soc. **2008**, *130*, 14916.

(70) Skurski, P.; Sobczyk, M.; Jakowski, A.; Simons, J. *Possible mechanisms for protecting N-Cα bonds in helical peptides from electron-capture (or transfer) dissociation* Int. J. Mass Spectrom. **2007**, *265*, 197.

(71) Basicity and nucleophilicity are of course different scales, but usually run parallel.

(72) Brodbelt, J. S.; Wysocki, V. H.; Cooks, R. G. *Thermochemical vs. Kinetic Control of Reactions in an Ion Trap Mass Spectrometer* Org. Mass Spectrom. **1988**, *23*, 54.

(73) Tsaprailis, G.; Nair, H.; Somogyi, A.; Wysocki, V. H.; Zhong, W.; Futrell, J. H.; Summerfield, S. G.; Gaskell, S. J. *Influence of Secondary Structure on the Fragmentation of Protonated Peptides* J. Am. Chem. Soc. **1999**, *121*, 5142.

Chapter 8.

Gas-phase conformations of small polyprolines and their fragment ions by IRMPD spectroscopy

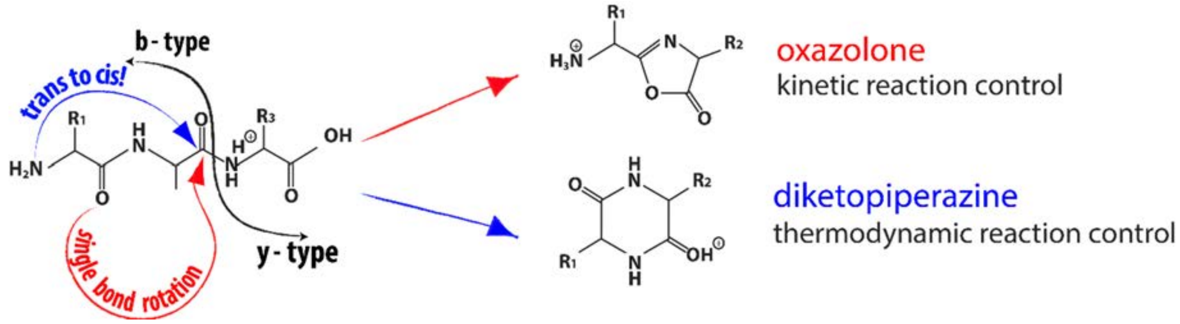
Infrared multiple-photon dissociation (IRMPD) spectroscopy has been used to examine the gas-phase conformations of a series of short protonated polyproline ions (Pro_3, Pro_4, Pro_5, Pro_6), their CID/IRMPD fragmentation pathways, and the associated fragment identities. Consistent with previous findings, and in combination with density functional theory (DFT) and MM/MD methods, a series of conformers for the protonated parent ions having their first peptide bond in the *cis* conformation has been identified. This conformation maximizes the solvation of the protonated N-terminus and stabilizes these compact globular-type conformations. This is in contrast to the PPI and PPII polyproline-type helices reported for larger polyproline peptides in solution. As well, this conformation leads to a unique fragmentation pattern upon collisional or multiple-photon activation. We report observation of the uncommon, but thermodynamically favoured, diketopiperazine-type b_2^+ fragment ion. Formation of b_2^+ ions along the diketopiperazine pathway is in line with a *cis* configuration of the first amide linkage in the protonated parent ion. Additionally, the parent ion conformations, fragmentation pathways, and proton affinities of the resulting fragments have been related to the observed proline-effect in CID mass spectra. *

---

*Adapted from: J.K.Martens, J.Grzetic, G.Berden and J.Oomens, *Gas-phase conformations of small polyprolines and their fragment ions by IRMPD spectroscopy*, Int. J. Mass. Spectrom. (2014)

8.1. Introduction

Mass spectrometry has developed into an indispensable tool in modern proteomics and biochemical analysis in general. Complex and sophisticated sequencing algorithms have been developed relating fragmentation patterns to parent ion sequence. However, much of the backbone fragmentation following non-standard pathways remains poorly modelled and the data resulting from these pathways is often not completely utilized[1], possibly resulting in reduced sequence information and less reliable identification. In order to more completely exploit the information contained in protein/peptide MS^n experiments, incorporation of a more comprehensive description of peptide fragmentation mechanisms, including not only models of non-standard fragmentation pathways, but also relative fragment-ion intensity information, is required[2].



Scheme 8.1. Different reaction mechanisms for formation of oxazolone and diketopiperazine b-type fragment-ions showing necessary backbone distortions for each reaction.

The fragmentation of protonated peptides can be described using the mobile proton model, in which a proton normally sequestered at the most basic site in the peptide can localize on amide nitrogen sites following activation[3-7]. This leads to cleavage of the resulting weakened $C(=O)-NH_2^+$ bond and is referred to as the b_n-y_m pathway[8]. As illustrated in Scheme 8.1 for b_2 formation, this most commonly occurs by nucleophilic attack of the first amide carbonyl oxygen on the carbon of the protonated amide leading to an oxazolone-type fragment[9-11]. Dissociation can also be accomplished with the N-terminal nitrogen acting as the nucleophile, forming a diketopiperazine-type fragment. Within this model, the oxazolone generating reaction requires *trans* conformation of the first peptide bond, while the diketopiperazine analog requires *cis* conformation. The oxazolone-type structure is common and thought to be kinetically favoured while the diketopiperazine-type structure is rare, but thermodynamically favoured.

While peptide bonds are nearly exclusively in the *trans* configuration, isomerization from *trans* to *cis* conformation of the peptide bond is in some cases thought to be feasible[12, 13, 14]. This can be attributed to the weakening and loss of double-bond character of the peptide bond after protonation of the amide nitrogen. This is especially possible when a basic residue traps the proton near to the site of isomerization and has been used to explain the observation of diketopiperazine b-type fragments from protonated histidine-containing peptides[14,15]. In other cases, such as in proline-containing peptides, the peptide bond may exist in the *cis* conformation in advance, making the formation of the diketopiperazine b-fragment more easily rationalized. *Cis* peptide bond conformation is commonly cited as occurring in only 0.03-0.05 % of cases as a result of both the difference in relative energy of the *cis/trans* conformers (~+10 kJ mol^{-1}) and the high barrier to isomerization (~85 kJ mol^{-1})[16,17]. However, for amide bonds involving a proline residue, the *cis/trans* conformers are nearly iso-energetic, being separated by only ~2 kJ mol^{-1} and connected by a barrier of approximately 55 kJ mol^{-1} [16,17].

Within the framework of the b_n-y_m fragmentation pathway, the proton is finally shared by the two neutral fragments of the protonated parent in a proton-bound-dimer (PBD)[18]. It is expected that the relative ratio of b_m/y_n ions should qualitatively correspond to the ratios of proton affinities (PAs) between the fragments[8]. As it is generally assumed that the y_n fragment (as a truncated peptide) should have a higher PA than the cyclic diketopiperazine, it is expected that upon dissociation of the PBD, the diketopiperazine fragment would often be lost as a neutral and remain unobserved. For oxazolone b-type fragments, having proton affinities typically on the order of the corresponding y-fragments, this is much less the case.

Proline is unique among the common amino acids due to its pyrrolidine ring and the associated rigidity it imparts on the peptide backbone when present. It has been suggested that *cis-trans* isomerization of the peptide bonds in a proline containing-protein is the rate-limiting step in the folding process[19]. Overly-represented fragmentation at Xxx-Pro amide linkages in peptides has been reported in many studies[20-25] and is commonly referred to as the *proline effect.*[26-30] Additionally, it has been noted that there is an uncommon over-expression of y-type fragment ions in the CID spectra of proline containing-peptides, as well as a dominant b_2^+ ion peak in the case of N-terminal Xxx-Pro occurrences[31]. These effects are often attributed to the high proton affinity of proline and the resulting C-terminal fragments (y-type) as well as the low threshold-energy related to cleavage N-terminal to proline[20]. CID experiments of Pro_8NH_2-H^+ have been used to demonstrate the near absence of b-type ions as well as unbalanced abundances of y-type ions, favouring those resulting from the loss of an even number of proline residues.[32] This fragmentation pattern was found to be dependent on an unblocked N-terminus and the absence of proton sequestering high-basicity residues. Additionally, it was noted that this unique pattern from repeating Pro-residues is not a result of sequential losses of diproline units, but rather the immediate loss of an even number of Pro residues from the protonated parent.

In solution, polyprolines commonly exist as either an all-*cis*, right-handed helix with 3.3 residues per turn (PPI) or an all-*trans* left-handed helix with 3.0 residues per turn (PPII)[33]. These helical structures are unique in that they completely lack hydrogen bonding in their peptide backbones. PPI is stabilized by a tight conformation, maximizing van der Waals interactions, while PPII is stabilized by solvent interactions with the outward facing carbonyl oxygens. PPI is additionally unique in that the macrodipole along the axis of the helix has its partial negative component towards the N-terminus and is thus stabilized by N-terminal protonation. This is in direct contrast to the more common alpha-helix (and others) where the helix is stabilized by an extensive network of backbone hydrogen bonds. In this case, protonation at the N-terminus interacts with the macrodipole in a destabilizing manner as a result of the consistent orientation of the amidic N-H groups towards the N-terminus. Ion mobility experiments and molecular dynamics simulations by Counterman and Clemmer have explored the conformational properties of a large range (over both size and charge state) of polyproline peptides[34]. Small polyprolines were found to mainly adopt conformations that solvate the protonated N-terminus, suggesting that for smaller systems the van der Waals stabilization may be insufficient to promote the formation of the helices. They reported a common structural motif involving *cis* conformation of the first peptide bond in the system, which leads to a bend in the backbone and strong hydrogen bonds between the protonated N-terminus and one or a few of the first carbonyl oxygens. In their simulations, conformations of the remaining amide bonds were found to be of less importance, where possibly a mixture of *cis/trans* linkages allowed for more compact gas-phase structures.

Infrared multiple photon dissociation (IRMPD) spectroscopy is a technique that provides direct conformational insight for gas-phase ions and has been applied in the past to identify the conformation of peptide fragment ions[35-38,15]. Here, we use IRMPD spectroscopy in compliment with computational methods to examine the conformations of a series of small polyprolines (3-6 residues) and the influence of the *cis/trans* amide bond conformation on their fragmentation pathways.

8.2. Experimental and Computational Methods

8.2.1. Infrared multiple photon dissociation spectroscopy

Infrared action spectra of protonated peptide and fragment ions were obtained using a previously described Fourier transform-ion cyclotron resonance (FT-ICR) mass spectrometer coupled to the FELIX infrared free electron laser[39]. Protonated molecular ions ($[M+H]^+$) ions were generated using electrospray ionization (ESI) from approximately millimolar peptide solutions in 50:50 acetonitrile:water, acidified with ~0.1% formic acid. Peptide samples were obtained from GeneCust (Luxemburg) at 95% purification and used as received. Pro_2 and Pro_6 were synthesized at Radboud University Nijmegen by Dr. Dennis Löwik and cyclo-ProPro was prepared at the University of Amsterdam by Dr. Jan van Maarseveen according to a literature protocol[40]. Following the electrospray process, $[M+H]^+$ ions were accumulated in a linear hexapole trap before being injected into the ICR cell. Peptide fragment ions were generated by nozzle-skimmer dissociation in the high-pressure region at the interface of the electrospray source and the hexapole trap/accumulation region, or directly in the ICR cell by irradiation with a 35 W continuous-wave CO_2 laser for a few hundred milliseconds (ULR-25; Universal Laser Systems, Scottsdale, USA).

Once trapped in the cell, all unwanted ions are ejected by a stored waveform inverse Fourier transform (SWIFT) excitation pulse. The remaining isolated ions are then irradiated by the tunable infrared radiation from FELIX, which arrives as 5 μs macropulses (5 or 10 Hz) of approximately 30-40 mJ with a bandwidth of ~0.4% of the center frequency. Resonant absorption leads to an increase in the internal energy resulting in unimolecular dissociation. This produces a frequency-dependent fragmentation signal which is recorded in the ICR cell using a typical excite/detect procedure. Relation of parent and fragment ion intensities in the observed mass spectral data in terms of the fragmentation yield (ΣI(fragment ions)/ΣI(parent + fragment ions)) generates the final infrared vibrational action spectrum. The yield at each IR point is obtained from three averaged mass spectra and is linearly corrected for laser power; the frequency is calibrated using a grating spectrometer. IRMPD spectra in the C-H, N-H and O-H stretching region (2700-3700 cm^{-1}) were obtained using an optical parametric oscillator (OPO, Laser Vision, Bellevue, USA). The Nd:YAG-pumped-OPO generates between 10 and 20 mJ per 5 ns pulse at 10 Hz in the 2700-3700 cm^{-1} range with a bandwidth of approximately 3 cm^{-1}. The IR frequency is calibrated with a wavemeter.

Collision induced dissociation (CID) mass spectral experiments were conducted using a Bruker AmaZon Speed ETD quadrupole ion trap. $[M+H]^+$ ions were generated using electrospray ionization (ESI) from approximately 10 micromolar peptide solutions in 50:50 acetonitrile:water acidified with 0.1% formic acid. After isolation in the trap, the $[M+H]^+$ ions of interest were subjected to CID conditions for 40 ms with amplitude parameters of approximately 0.5-1.0 V.

8.2.2. Computational chemistry

Even small peptides, such as those considered here, tend to have relatively complex conformational landscapes. As a result, manual definition of input geometries using chemical intuition alone rapidly becomes difficult when considering peptides of sizes larger than 2-3 residues. In order to more completely explore the potential energy surface of these peptides, we have therefore employed a molecular mechanics/molecular dynamics (MM/MD) approach using AMBER 12[41], approximately following a procedure described in more detail previously[42]. Initial guess geometries were first optimized in the Gaussian09 package[43] at the B3LYP/6-31++G(d,p) level of theory. Charges from these initial results were used for parameterization of the nonstandard terminations of the peptides in their predicted non-zwitterionic gas-phase conformations in the antechamber program. After minimization within AMBER, a simulated annealing procedure up to 1000 K was used. With this temperature range it was found that even relatively large barriers, such as *cis/trans* isomerization of the peptide bond, were adequately overcome during the simulations of these polyproline peptides. A 1 fs step size was used with coordinates being saved

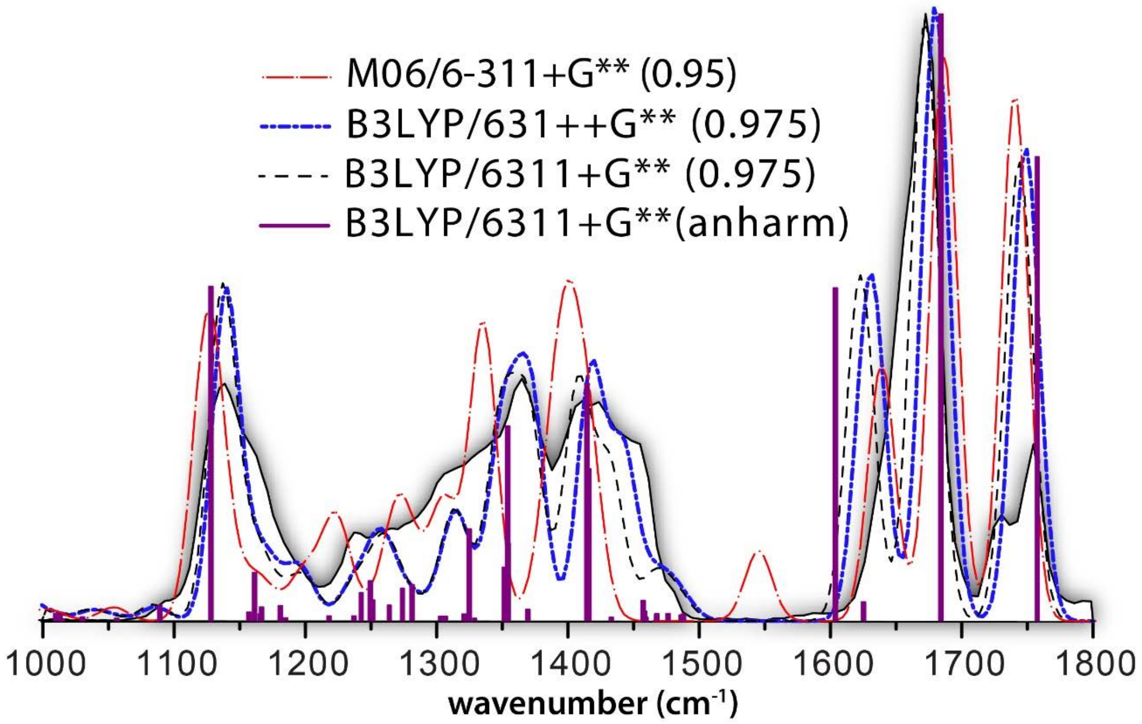


Figure 8.1. Comparison of calculated vibrational spectra and IRMPD spectrum in the 1000-1800 cm^{-1} fingerprint region

periodically throughout the simulation for individual optimization. The resulting 500 structures were grouped based on structural similarity using appropriate rms criteria (leading to approximately 30 structures for each peptide size). Next, these structures were each optimized at the B3LYP/6-31++G(d,p) level followed by vibrational analysis within the rigid-rotor harmonic oscillator model. Vibrations were scaled by 0.975 in the fingerprint region (800-2000 cm^{-1}), while C-H, N-H and O-H stretches in the 2700-3700 cm^{-1} region were scaled by 0.955. Calculated stick spectra are broadened using a Gaussian function with a FWHM of 25 cm^{-1} for comparison with experiment. For Pro_3H^+, optimizations on all geometries resulting from the MM/MD procedure were also completed at the B3LYP/6-311+G(d,p) level, and MP2(full)/6-311+G(2d,2p) electronic energies were obtained for optimized geometries at both DFT levels. Finally, harmonic frequencies using the M06 functional as well as anharmonic frequencies at the B3LYP/6-311+G(d,p) level were obtained for the lowest energy conformation of Pro_3H^+. Energies at all three levels of theory were found to be largely consistent. As well, M06 frequencies and anharmonic frequencies were not found to offer improvement over modelling at the B3LYP/6-31++G(d,p) level. These additional calculations are summarized in the supporting information file in Table 8.1 and Figure 8.1. Proton affinities (PAs) and gas-phase basicities (GBs) were calculated using thermodynamic quantities for the protonated and neutral peptides and fragment species at the B3LYP/6-31++G(d,p) level.

Table 8.1. Comparison of Calculated Energies - Pro_3H^+ (kJ mol^{-1})

	B3LYP/6-31++G(d,p)			MP2/6-311+G(2d,2p)//B3LYP/6-31++G(d,p)		
	E_{rel}	**H_{rel}**	**$G_{rel,298K}$**	**E_{rel}**	**H_{rel}**	**$G_{rel,298K}$**
pro₃H_c0	7.2	6.7	5.8	13.3	12.8	12.0
pro₃H_c1	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
pro₃H_c2	39.2	36.5	35.9	41.9	39.2	38.7
pro₃H_c3	9.4	9.2	12.7	3.6	3.5	7.0
pro₃H_c4	7.0	4.6	0.9	13.4	11.0	7.3
pro₃H_c6	9.4	7.1	2.8	17.6	15.3	11.0

8.3. Results and Discussion

Figures 8.2-8.6 show the IRMPD spectra of Pro_3H^+, Pro_4H^+, Pro_5H^+ and Pro_6H^+ obtained in the 800-1800 cm^{-1} spectral region as well as the 2700-3700 cm^{-1} region of C-H, N-H and O-H stretches (measurements in the 2700-3700 cm^{-1} region were not obtained for Pro_6H^+). Figure 8.7 shows a direct comparison of these experimental spectra. IRMPD spectra of the b_2^+, y_2^+ (from the Pro_3H^+ precursor) and y_3^+ (from the Pro_4H^+ precursor) are presented in Figures 8.9, 8.10 and 8.2, respectively. CID mass spectra of the protonated parent ions, Pro_2H^+, Pro_3H^+, Pro_4H^+, Pro_5H^+, Pro_6H^+ and Pro_7H^+, are presented in Figure 8.7. The MM/MD/DFT procedure described above was used to obtain around 30 optimized geometries and vibrational spectra for each of the Pro_3H^+, Pro_4H^+, Pro_5H^+, $Pro_6H^+$and y_2^+ species. The peptide bond dihedral angles, relative electronic energies, enthalpies and Gibbs free energies (298 K) are presented in Tables 8.2-8.6 for the six lowest energy structures found. Dihedral angles are qualitatively labelled as being in either cis or trans conformation, however some variation from zero and 180 degrees does exist in most cases and precise values are given in the supporting information (Tables 8.2-8.6). Calculated conformations are organized with the addition of an index "_c#" after the name of the peptide.

Table 8.2. y_2^+ (kJ mol^{-1})

B3LYP/6-31++G(d,p)	$\Omega_{peptide}$	E_{rel}	H_{rel}	$G_{rel,298K}$
pro₂H_c2	-18	5.6	6.9	6.1
pro₂H_c6	-8	3.0	3.1	1.8
pro₂H_c19	170	0.0	0.0	0.0
pro₂H_c24	-3	14.7	14.1	11.9
pro₂H_c35	177	13.1	12.9	12.1
pro₂H_c40	0	44.2	43.6	42.6

Table 8.3. Pro$_3$H$^+$ (kJ mol^{-1})

B3LYP/6-31++G(d,p)	$\Omega_{peptide}$	E_{rel}	H_{rel}	$G_{rel,298K}$
pro$_3$H_c0	-17,174	7.2	6.7	5.8
pro$_3$H_c1	-8,-13	0.0	0.0	0.0
pro$_3$H_c2	170,9	39.2	36.5	35.9
pro$_3$H_c3	-12,173	9.4	9.2	12.7
pro$_3$H_c4	164,173	7.0	4.6	0.9
pro$_3$H_c6	166,172	9.4	7.1	2.8

Table 8.4. Pro$_4$H$^+$ (kJ mol^{-1})

B3LYP/6-31++G(d,p)	$\Omega_{peptide}$	E_{rel}	H_{rel}	$G_{rel,298K}$
pro$_4$H_c0	-13,-14,172	17.4	17.1	12.0
pro$_4$H_c2	-24,123,-6	23.7	21.6	24.1
pro$_4$H_c3	-11,170,-175	0.0	0.0	0.0
pro$_4$H_c5	-12,178,-7	11.6	11.5	8.7
pro$_4$H_c8	-8,-17,-10	15.2	14.6	10.3
pro$_4$H_c17	-9,-15,174	19.0	18.8	14.8

Table 8.5. Pro$_5$H$^+$ (kJ mol^{-1})

B3LYP/6-31++G(d,p)	$\Omega_{peptide}$	E_{rel}	H_{rel}	$G_{rel,298K}$
pro$_5$H_c0	-3,175,2,-170	0.0	0.0	0.0
pro$_5$H_c2	163,168,168,172	65.7	65.7	42.3
pro$_5$H_c3	0,179,-167,179	28.5	29.2	28.0
pro$_5$H_c4	-12,-13,167,173	40.0	41.8	25.6
pro$_5$H_c5	-8,-19,-15,172	35.9	38.1	23.2
pro$_5$H_c7	-10,171,-31,178	27.3	29.0	17.6

Table 8.6. Pro$_6$H$^+$ (kJ mol^{-1})

B3LYP/6-31++G(d,p)	$\Omega_{peptide}$	E_{rel}	H_{rel}	$G_{rel,298K}$
pro$_6$H_c8	-13,176,6,-162,172	0.0	0.0	0.0
pro$_6$H_c15	-21,176,7,-170,-178	5.9	5.2	2.0
pro$_6$H_c9	-14,-177,164,164,175	23.9	23.1	9.0
pro$_6$H_c32	-7,-20,-16,-10,-180	22.1	21.9	11.3
pro$_6$H_c30	-14,-22,168,2,-7	32.4	32.1	21.8
pro$_6$H_c31	142,171,165,166,-2	82.4	80.8	64.5

8.3.1. Structures of Protonated Pro_3, Pro_4, Pro_5 and Pro_6

Structure of Pro_3H^+. The lowest energy calculated conformation of Pro_3H^+, pro_3H_c1, has both peptide bonds in the *cis* conformation and the protonated N-terminus interacting with both carbonyls of the first and second residues, as well as the C-terminal carboxyl group. The close agreement between the calculated spectrum of this ion and the measured IRMPD spectrum leads us to assign this conformation to that of Pro_3H_c1 under our experimental conditions (see SI for comparison with other low-energy structures). Figure 8.2 also contains the IRMPD spectrum of the y_3^+ ion generated from Pro_4H^+. As expected, this spectrum closely matches that of Pro_3H^+ and leads to its assignment as a protonated truncated peptide. The double feature of the carboxyl stretch between 1700-1800 cm^{-1} in both the spectrum of Pro_3H^+ and that of the y_3^+ fragment, is unaccounted for by the single band in the calculated structure of Pro_3H_c1 and likely indicates some contribution of a second low-energy conformer, possibly Pro_3H_c3 with a *cis/trans* conformation, see Figure 8.3. The calculated spectrum in the 2900-3700 cm^{-1} region contains C-H stretches below 3000 cm^{-1} and N-H stretches between 3100 and 3300 cm^{-1}. These bands are, however, unobserved experimentally because of the relatively low pulse energy from the OPO below ~3000 cm^{-1} and the likely high degree of anharmonicity associated with the strongly hydrogen-bonded amine N-H stretches[44-46]. The unbound carboxyl CO-H stretch at ~3590 cm^{-1} corresponds very well to the experimentally measured band in this region.

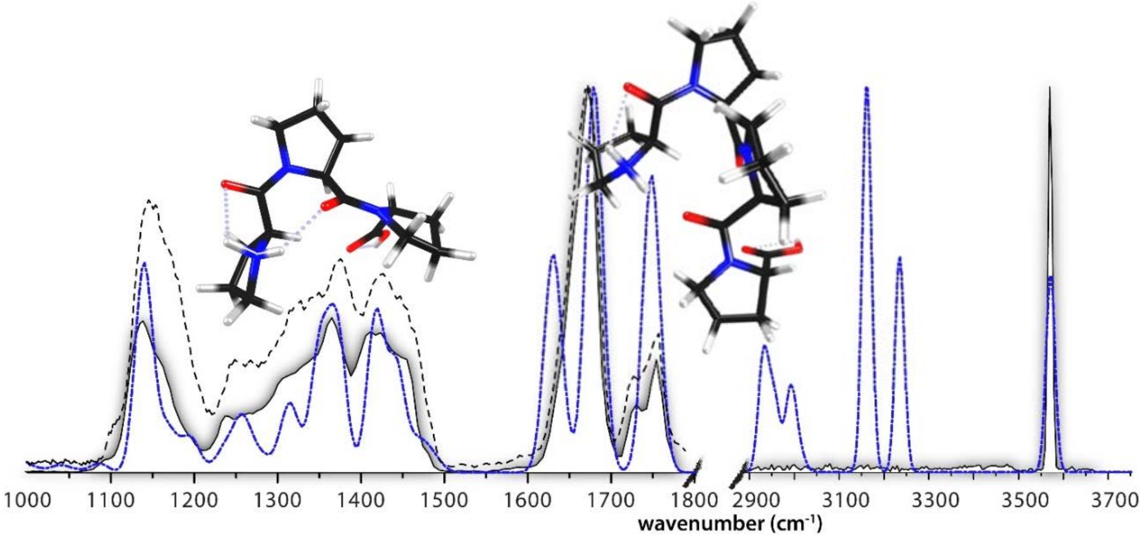


Figure 8.2. IRMPD spectrum of Pro_3H^+ shown with the calculated vibrational spectrum of the lowest energy structure, pro_3H_c1, in blue (structure inlayed) with cis/cis conformation of the amide bonds in the peptide backbone. The experimental spectrum of the y_3^+ fragment ion produced from CID of Pro_4H^+ is shown in black dashed line.

Structure of Pro_4H^+. Figure 8.4 contains the calculated spectra of the two lowest-energy identified Pro_4H^+ conformers, pro_4H_c3 and pro_4H_c5, overlaid on the experimental IRMPD spectrum. While the second lowest-energy conformer, pro_4H_c5, is calculated to be 8.7 kJ

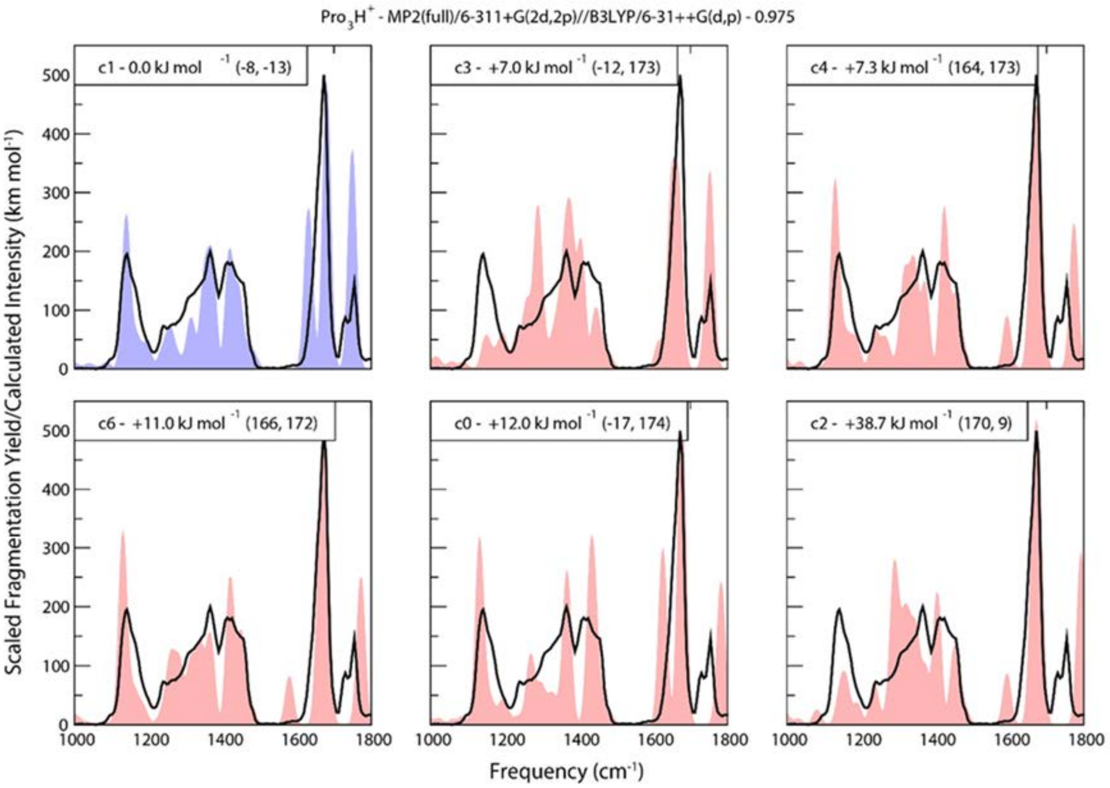


Figure 8.3. Lowest-energy calculated structures with spectra for Pro_3H^+. Relative free energies (298 K) are given in each pane with the dihedral angles (Ω) listed starting from the N-terminus (note, MP2 electronic energies are used in calculation of free energies solely in this plot for Pro_3H^+).

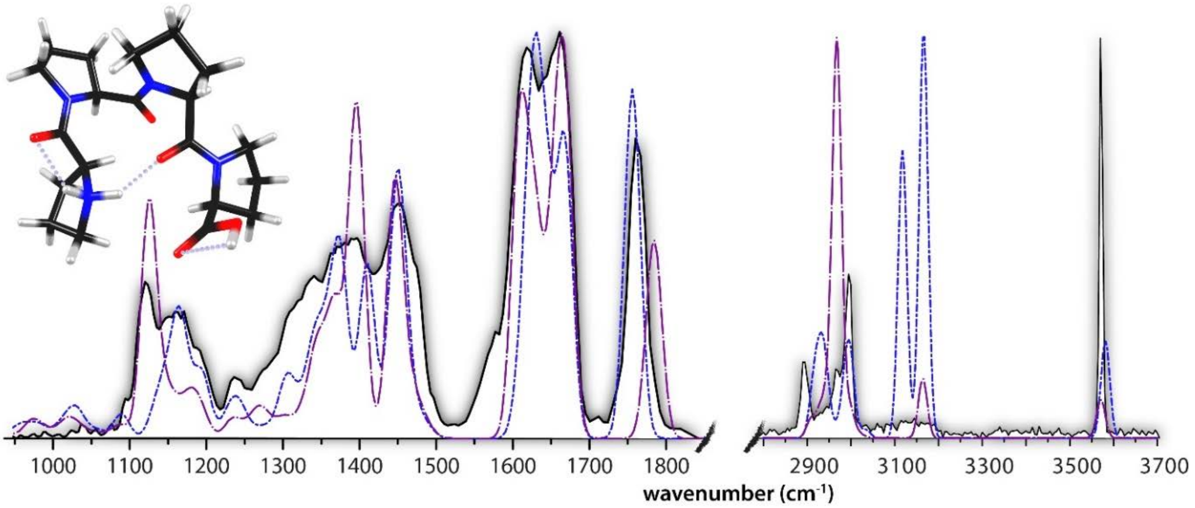


Figure 8.4. IRMPD spectrum of Pro_4H^+ shown with the computed vibrational spectrum of the lowest-energy structure (pro_4H_c3, cis/trans/trans, structure inlayed left) in blue and the second lowest energy structure (pro_4H_c5, 8.7 kJ mol^{-1}, cis/trans/cis, structure inlayed right) in dashed purple.

mol^{-1} higher in free energy, the lowest energy conformer, pro_4H_c3, alone appears to be unable to account for the C-OH bending feature observed around 1100 – 1200 cm^{-1} and possibly the region of highest intensity immediately to the red of 1400 cm^{-1}. The second lowest energy conformation accounts for these regions nicely and we suspect a mixture of these two conformers to co-exist. The differences in the C-OH bends is likely a result of the long-range weak hydrogen bond involving the C-terminus in pro$_4$H_c5 that does not exist in the structure of pro$_4$H_c3, both shown in Figure 8.4.

In the 3 μm spectral range, a small shoulder-peak in the C-H stretching band near 2960 cm^{-1} matches well to the intense peak in the spectrum of the second lowest-energy structure, providing additional support for a fractional presence of this structure. The very weak broadly distributed intensity in the 3100 cm^{-1} region is better accounted for by the lowest-energy conformation, as well as the C-H stretches to the red of 2900 cm^{-1}, although here, the lowest-energy conformation has a small blue-shift in comparison with the experimental band. The free CO-H stretch at ~3590 cm^{-1} is in good agreement with the calculated bands of both structures.

Structure of Pro$_5$H$^+$. Pro$_5$H$^+$ has a somewhat different calculated lowest-energy conformation (pro$_5$H_c0) in comparison with Pro$_3$H$^+$ and Pro$_4$H$^+$, as illustrated in Figures 8.5 and 8.7. The main difference is the conformation of the C-terminal CO-H group, where the proton is shared between the C-terminus and the C=O of the fourth residue. In addition, the N-terminus is interacting with the C=O of the C-terminus and the C=O of the first residue. In this bound conformation, the CO-H stretch is, as expected, red-shifted and broadened and thus unobserved at 3570 cm^{-1}, where it was observed for Pro$_3$H$^+$ and Pro$_4$H$^+$. There is, however, a very small band observed in the experimental spectrum at 3570 cm^{-1}, attributed to a contribution from a minor population of the second lowest-energy structure (pro$_5$H_c7, 17.6 kJ mol^{-1} higher in free energy). The broadly distributed intensity centred at 2700 cm^{-1} corresponds well to the CO-H stretch of the lowest-energy conformation. The C-H stretch bands of the second lowest energy conformation possibly contribute a slight shoulder to the more intense C-H stretching band of the lowest-energy conformation just to the red of 3000 cm^{-1}.

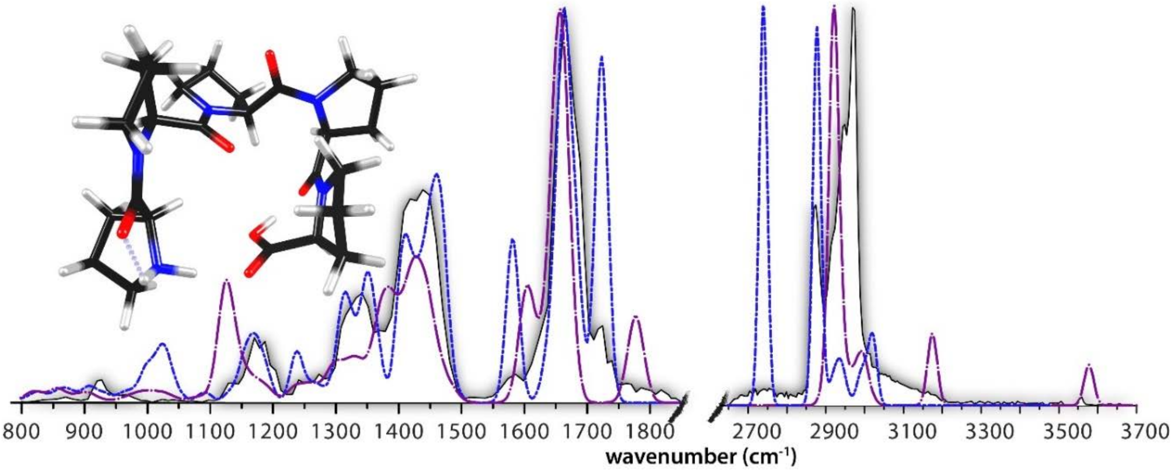


Figure 8.5. IRMPD spectrum of Pro$_5$H$^+$ shown with calculated vibrational spectrum of the lowest energy calculated structure in blue (pro$_5$H_c0, cis/trans/cis/trans, structure inlayed) and the second lowest energy calculated structure (pro$_5$H_c7, 17.6 kJ mol^{-1}, cis/trans/cis/trans) in dashed purple.

There is relatively clear disagreement between the experimental spectrum and that computed for the second lowest-energy conformer in the fingerprint region, further suggesting the contribution from ions having the second lowest-energy conformation to be very limited. The line positions of the calculated amide I vibrations, carboxyl C=O stretch just above 1700 cm^{-1}, and the N-H scissoring mode at 1580 cm^{-1} for pro$_5$H_c0 match nicely with the IRMPD spectrum. Both experimental and calculated spectra above 1550 cm^{-1} are unique from Pro$_3$H$^+$ and Pro$_4$H$^+$ and clearly contrast the structural differences. The calculated spectrum of pro$_5$H_c7, while likely representing a small percentage of the ion population, is again seen to contribute only slightly to the observed spectrum. Possibly, a small shoulder in the experimental COH bend at 1130 cm^{-1} as well as a non-zero baseline around 1775 cm^{-1} is observed from ions having this conformation.

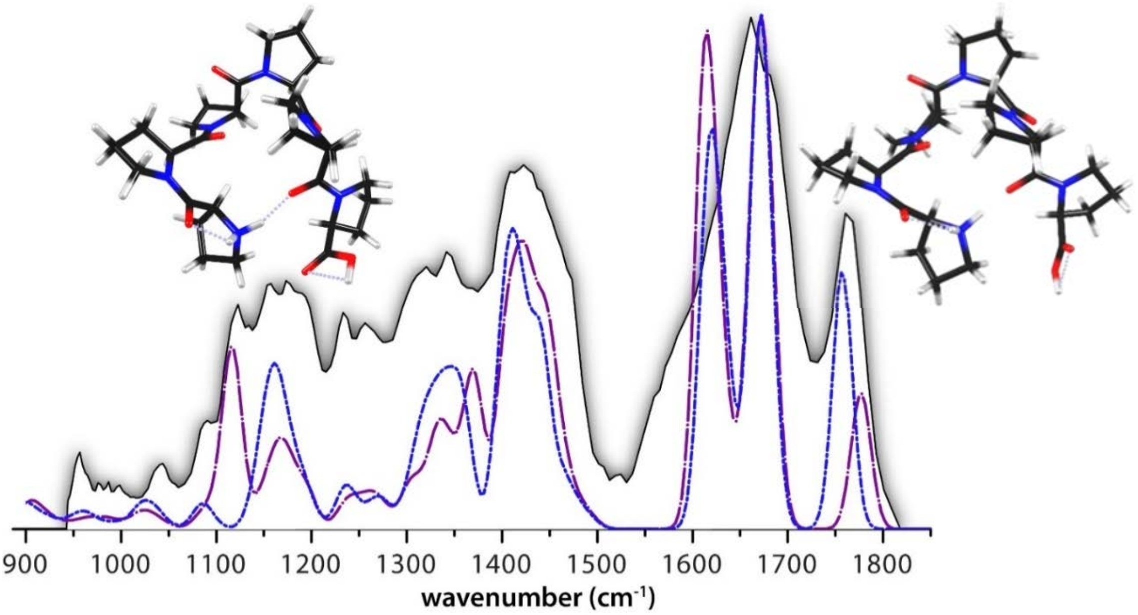


Figure 8.6. IRMPD spectrum of Pro$_6$H$^+$ shown with the calculated spectra of the lowest energy structure in blue (pro$_6$H_c8, cis/trans/cis/trans/trans, structure inlayed left) and the second lowest energy structure (pro$_6$H_c15, 2.0 kJ mol^{-1}, cis/trans/cis/trans/trans, structure inlayed right) in dashed purple.

Structure of Pro$_6$H$^+$. The IRMPD spectrum of Pro$_6$H$^+$, presented in Figure 8.6 matches reasonably well with the calculated spectra of the identified low energy structures, pro$_6$H_c8 and pro$_6$H_c15 (+2.0 kJ mol^{-1}). The calculated spectrum of pro$_6$H_c8, with the N-terminus hydrogen-bonded to the carbonyl group of the fifth residue, matches well in the 1600-1800 cm^{-1} amide I region. Perhaps consideration of the nearly iso-energetic second lowest energy conformation pro$_6$H_c15 accounts for the peak near 1125 cm^{-1}. In this conformer, the C-terminus is further away from, and thus less affected by, interaction with the charged N-terminus.

8.3.2. Spectral differences for Pro_3H^+, Pro_4H^+, Pro_5H^+ and Pro_6H^+ related to conformations

The IRMPD spectra of Pro_3H^+ and Pro_4H^+ are fairly similar with only a few distinguishing features, as illustrated in Figure 8.7. Indeed, the calculated structures which have been attributed to the experimentally observed spectra are similar. In both structures, the protons of the N-terminus are each hydrogen-bonded to adjacent carbonyl oxygens. In Pro_3H^+, these are the C=O groups of the first and second residues, but for Pro_4H^+ the first and third carbonyl oxygens are bound to the N-terminus, leaving the second amide carbonyl free. As a result, the partly overlapping C=O stretching bands of the three carbonyl groups produce the double-featured appearance of the amide I peak. The NH_2-scissoring mode of the protonated N-terminus is overlapping with this peak.

In the calculated spectrum of Pro_3H^+, the carbonyl stretch of the first residue is observed at approximately at 1680 cm^{-1}, while two overlapping modes at approximately 1630 cm^{-1} are both coupled motions of the second carbonyl stretching and the protonated N-terminus scissoring.

In contrast to Pro_3H^+ and Pro_4H^+, the amide I band for Pro_5H^+ is nearly overlapping with the red-shifted carboxyl C=O stretch, likely as a result of the different bonding environment of the C-terminus. The Pro_4H^+ COH bending peak just to the blue of 1100 cm^{-1} has a double

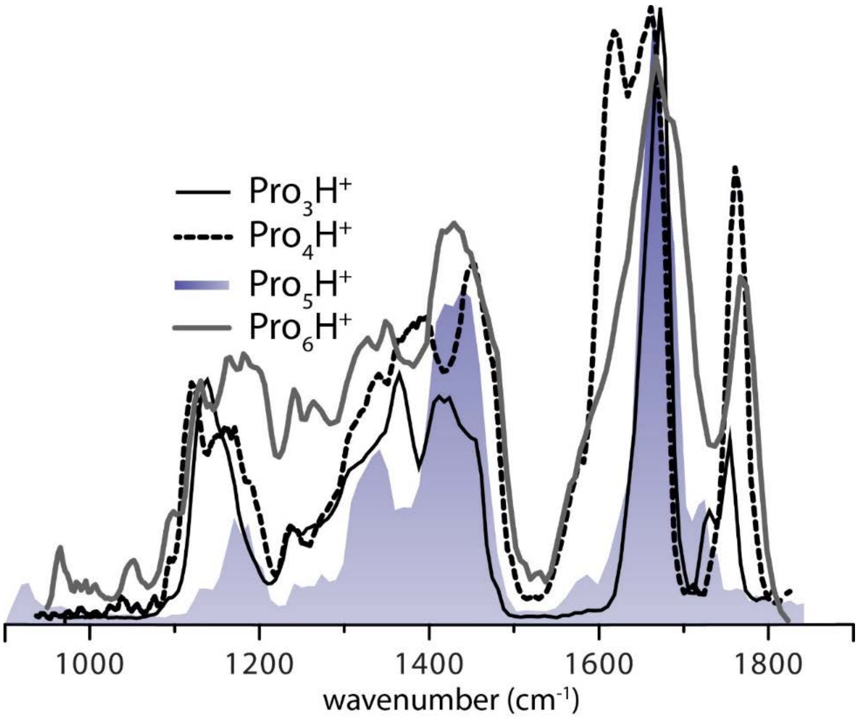


Figure 8.7. Overlay of IRMPD spectra for Pro_3H^+, Pro_4H^+, Pro_5H^+ and Pro_6H^+. Pro_5H^+ is markedly different in the 1700-1800 cm^{-1} region due to the alternative conformation of the carboxyl group in comparison with the other oligo-proline peptides.

feature in the IRMPD spectrum likely due to the contribution of the second lowest energy structure (illustrated in Figure 8.4). This is not observed for Pro_3H^+. Pro_5H^+ does not have a peak in this region, as expected, due to hydrogen-bonding of the CO-H with the carbonyl oxygen of the fourth residue. The peak just below 1200 cm^{-1} in the spectrum of Pro_5H^+ is mainly from coupled $-CH_2-$ ring vibrations, likely corresponding to the small shoulder of the C-OH bending peak for Pro_4H^+, and possibly Pro_3H^+.

The spectrum of Pro_6H^+ is again similar those observed for Pro_3H^+ and Pro_4H^+, suggesting a general structural pattern for polyprolines, with Pro_5H^+ as somewhat of an exception in the series. Generally speaking, the IRMPD spectra of these four protonated peptides exhibit similar overall features, while the details nicely distinguish the structural differences in the conformations.

8.3.3. Fragments and Dissociation pathways of Pro_3H^+

Structure of Pro_3H^+ b_2^+. IR spectroscopy of the b_2^+ fragment generated from Pro_3H^+ clearly shows that this fragment ion is of diketopiperazine type (see Figure 8.9). This can be inferred from the close spectral match with the spectrum calculated for the lowest-energy structure, which is of diketopiperazine type, and with the experimental IRMPD spectrum generated from the synthetic protonated cyclo-ProPro (see Figure 8.8). Additionally, the b_2^+ experimental spectrum shows no sign of a band near 1850 cm^{-1}, where the oxazolone C=O stretch would typically be found[35]; the calculated diketopiperazine amide C=O stretch at 1730 cm^{-1} matches favorably with experiment. Comparison with other calculated diketopiperazine structures, protonated at different sites and with different ring conformations (not shown), confirmed the assignment of the oxygen-protonated structure

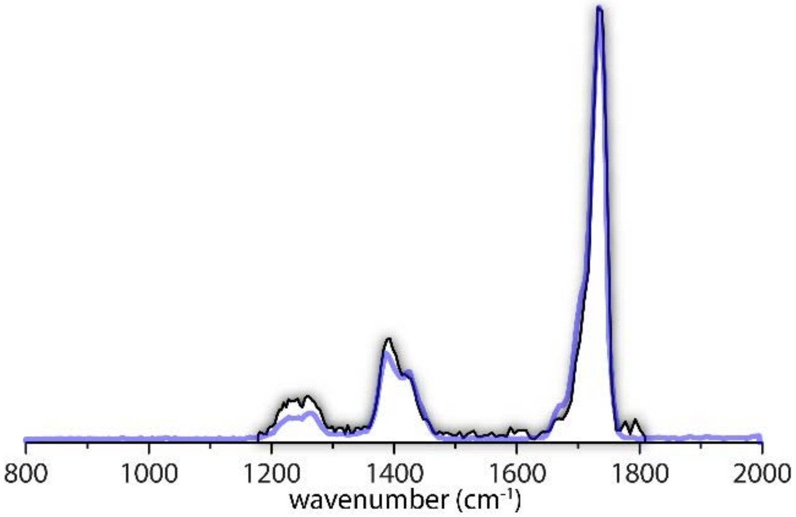


Figure 8.8. Comparison of the experimental IRMPD spectra of the protonated cyclo Pro-Pro reference compound (blue) with that of the b_2 fragment ion from protonated Pro_3 (black). All bands are seen to be reproduced exactly; the slight deviation in relative intensities is probably due to a slow decline of laser power during the scan of the reference compound, which is not accounted for in the power profile used for normalization.

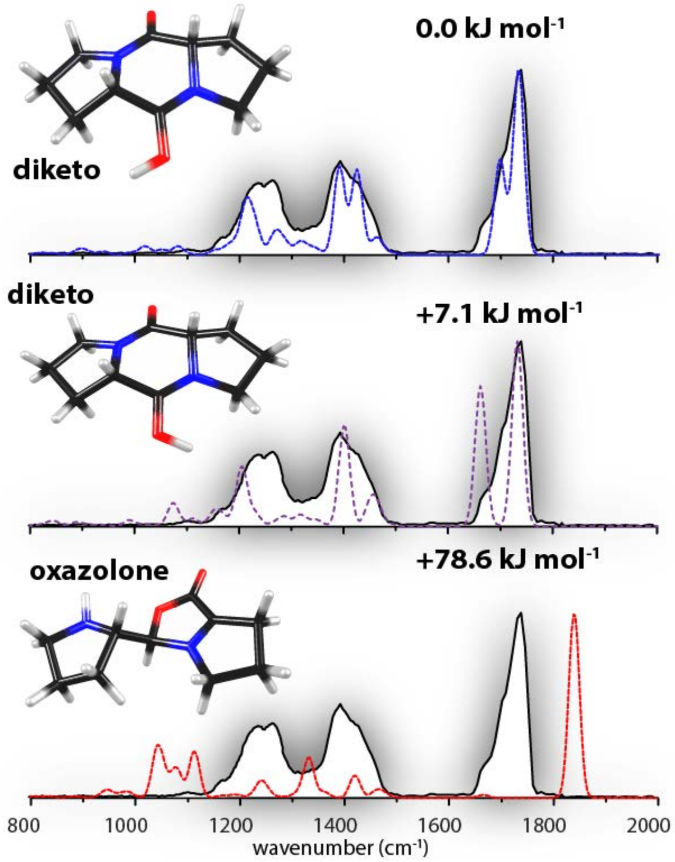


Figure 8.9. IRMPD spectrum of the b_2^+ fragment ion from Pro_3H^+ shown with the calculated vibrational spectrum of the lowest-energy diketopiperazine structure in blue and the second lowest diketopiperazine structure in purple. The lowest-energy calculated oxazolone fragment structure is shown in the bottom panel in dashed red line.

shown here. A shoulder in the experimental band to the red of 1700 cm^{-1} suggests that a portion of the sampled ions exists as the calculated higher-energy rotamer of the lowest-energy structure shown in the middle panel of Figure 8.9. Although only separated by 7.1 kJ mol^{-1}, interconversion between these two structures was computed to have a barrier of approximately 35 kJ mol^{-1}.

Structure of the y_2^+ fragment of Pro_3H^+. Figure 8.10 contains the IRMPD spectrum of the y_2^+ fragment ion derived from Pro_3H^+ as well as the calculated spectra for the lowest and, nearly-degenerate, second lowest-energy structures of the truncated peptide Pro_2H^+. Interestingly, these two conformations differ only by the cis/trans orientation of the peptide bond, with the trans conformation being 1.8 kJ mol^{-1} lower in free energy. This change in peptide bond conformation results in the first carbonyl group interacting with the other proton of the protonated N-terminus, and being closer to the $-CH_2-$ group of the proline ring instead of the proton attached to the α-carbon of the second residue. Such minor structural differences, predictably, result in very similar vibrational spectra, with the main differences being in the experimentally unobserved C-H stretches below 3000 cm^{-1} and the 1200-1300 cm^{-1} region of the fingerprint spectra. While the experimental spectrum matches qualitatively better to the lowest-energy trans conformation (blue), the contrast with the calculated spectrum of the second lowest-energy conformation is insufficient to assign one structure to the experimentally measured ion population. Additionally, consideration of the calculated free energies would suggest a mixture of these two conformations at room temperature.

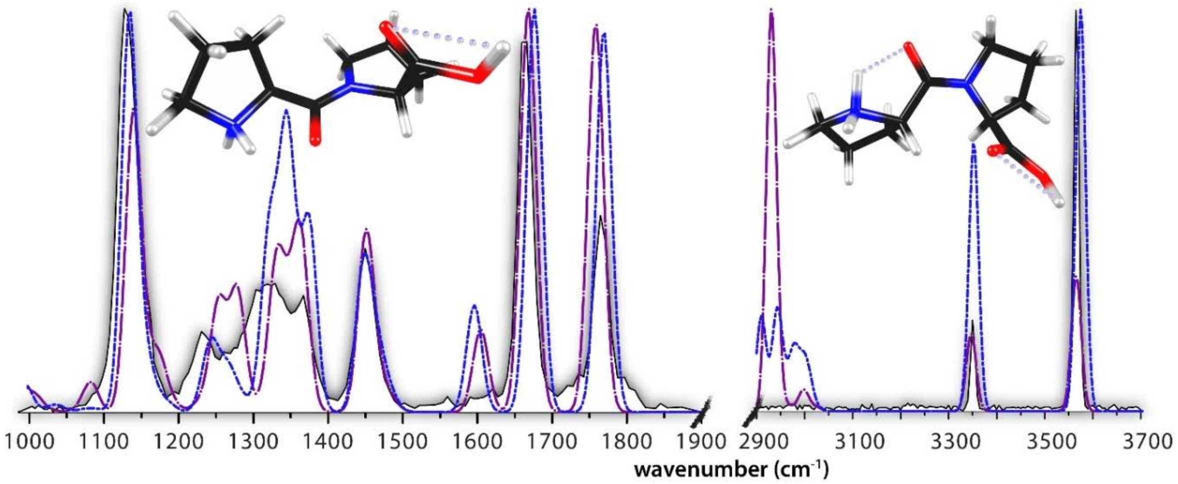


Figure 8.10. - IRMPD spectrum of y_2^+ (Pro_2H^+-equivalent ion) shown with calculated vibrational spectrum of the lowest energy calculated structure in blue (inlayed on the left) and the second lowest energy calculated structure (inlayed on the right) in dashed purple.

Table 8.7. CID Fragments Formed from $Pro_{2\text{-}7}H^+$ (%)

	Pro_2H^+	Pro_3H^+	Pro_4H^+	Pro_5H^+	Pro_6H^+	Pro_7H^+
m/z 70	29.0	-	-	-	-	-
y_1^+	71.0	5.5	2.6	-	-	-
m/z 126	-	11.9	1.4	-	-	-
a_2^+	-	1.8	0.3	0.4	-	-
b_2^+	**-**	**21.0**	**3.3**	**5.5**	**0.2**	**0.4**
y_2^+	-	59.7	58.8	16.6	14.1	3.1
y_3^+	-	-	33.7	67.1	7.3	27.3
y_4^+	-	-	-	10.5	76.3	3.8
y_5^+	-	-	-	-	2.0	65.3
Σy_n^+/total frag. Int.	**71.0**	**65.2**	**95.1**	**94.2**	**99.8**	**99.6**

CID of $Pro_{2\text{-}7}H^+$. Structural insight from the IRMPD results discussed above, as well as calculated thermochemical properties obtained for the matching conformations, provides valuable information for interpreting the observed CID fragmentation patterns of these polyprolines. Table 8.7 contains a list of the significant fragments observed after CID of the protonated Pro_2H^+ – Pro_7H^+ parent ions presented in Figure 8.11. For Pro_nH^+ (n=4-7) the most intense fragments are the y_{n-2}^+ ions[32] and very little intensity was observed for any of the b_m^+ ions and for the y_1^+ fragment. Pro_3H^+ is markedly different in its fragmentation pattern, having an intense b_2^+ signal. Pro_3H^+ has been previously reported to form mainly y- and b-type fragments (67.6% - y_2^+, 29.7% - b_2^+) and additionally at higher energies an ion at m/z 126[22]. This ion was proposed to be formed from both the b_2^+ ion and the $[M+H]^+$ precursor. These results are in close agreement with the presently reported CID results, with only small amounts of the a_2^+ and y_1^+ ions being observed here in addition.

Fragmentation along the b_m-y_n pathway reportedly results in a proton-bound-dimer of the complementary neutral fragments[4]. Upon dissociation of the dimer, the fragment retaining the proton will be the one observed in the mass spectrum and its counterpart will be lost as a neutral. Presumably, the fragment with the higher PA is expected to predominantly retain the proton. Table 8.8 lists the PAs and GBs of Pro_3 and its observed fragments. As y_n-type fragments are thought to correspond to truncated peptides (see also Figure 8.10), the calculated peptide PAs are also taken as the PAs of the corresponding neutral y_n-type fragments. As illustrated in Table 8.8, the PA of the diketopiperazine b-type fragment is lower than both the oxazolone b-fragment and any of the y_n-fragments considered (themselves comparable). This is consistent with the near absence of b-ions in the fragmentation patterns observed for Pro_4H^+, Pro_5H^+, Pro_6H^+ and Pro_7H^+ assuming that the b_2 fragments have diketopiperazine and not oxazolone structure. This assumption appears justified by the fact that all of the low-energy conformations for Pro_3H^+, Pro_4H^+, Pro_5H^+ and Pro_6H^+ identified here from the IRMPD spectra have *cis* conformation of their first peptide bond. The values in Table 8.8, however, do not explain the observed intensity ratios of the b_2^+ and y_1^+ fragments for Pro_3H^+, although for this combination, the PA-difference between b_2 and y_{n-2} fragments is by far the smallest in the series considered.

Clearly the fragmentation patterns observed here cannot be simply described using thermodynamic arguments and considerations of the internal energy distribution of the ions as well as kinetic factors need to be considered.

Table 8.8. Proton Affinities and Gas-Phase Basicities of Pro_3 and Associated Fragment Ions (kJ mol^{-1})

B3LYP/6-31++G(d,p)	PA	GB	Fragment Label
Pro_3	1028	1003	y_3^+
Pro_2	996	972	y_2^+
Proline	944	922	y_1^+
b_2-oxazolone	1025	1009	b_2^+
b_2-diketo	888	868	b_2^+

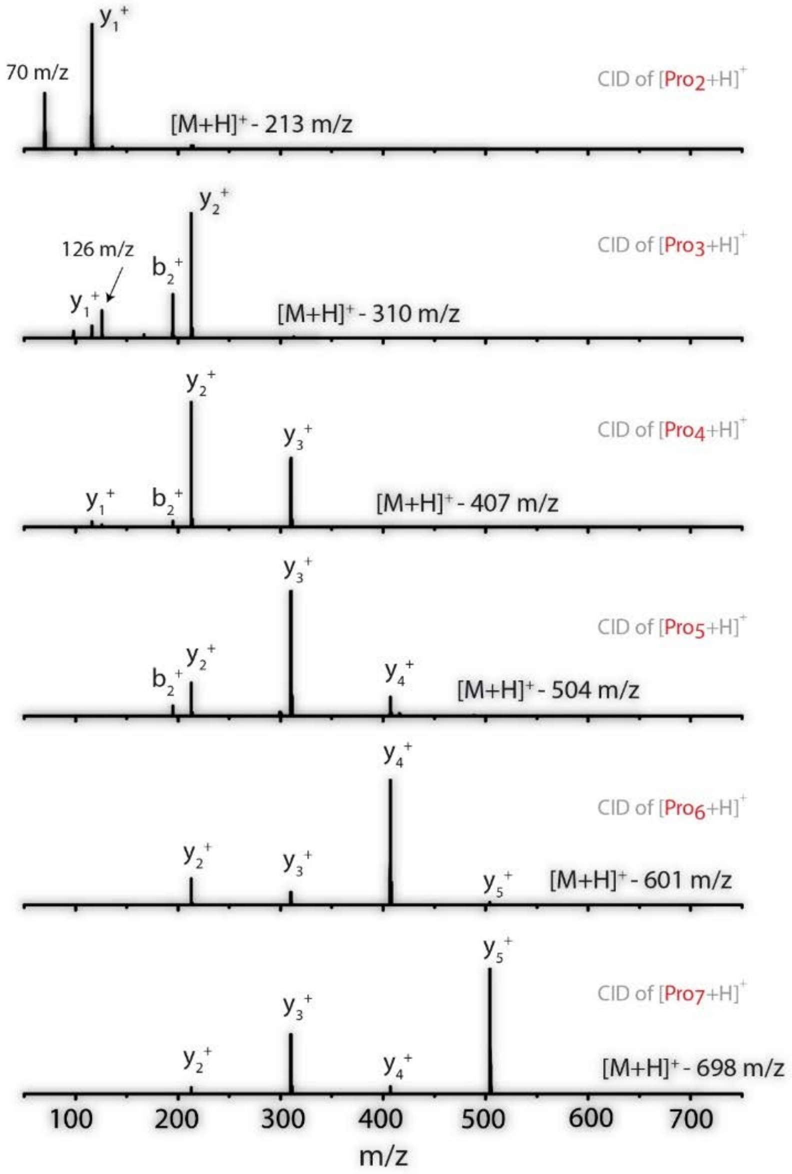


Figure 8.11. CID Mass spectra for Pro_nH^+ (n=2-7) precursor ions. Labelled ions: y_5^+ - 504 m/z , y_4^+ - 407 m/z, b_4^+ - 389, y_3^+ - 310 m/z, y_2^+ - 213 m/z, b_2^+ - 195 m/z, 126 m/z (see text for description), y_1^+ - 116 m/z.

8.4. Conclusions

In summary, the results presented here indicate that the b_2 fragment of Pro_3H^+ is characteristic of the thermodynamically-favoured diketopiperazine structure, showing no band higher than 1800 cm^{-1}, where an oxazolone carbonyl stretch would typically be found. This leads us to assign this ion to be one of only a few cases where a b-type fragment does not form on the so-called oxazolone pathway. Additionally, we observed collisional activation of the larger polyprolines to produce mainly *y*-type fragments. Computations show that *y*-ions of increasing size are increasingly likely to retain the proton in the dissociation process, particularly when the complementary N-terminal b_2-fragment possesses a diketopiperazine structure. This is due to the substantially lower proton affinity of the diketopiperazine structure as compared to the oxazolone b_2-structure.

Further evidence for the formation of diketopiperazine b_2-fragments is provided by IR spectroscopy, establishing that the first peptide bond in the protonated polyproline precursor ions is generally in the *cis*-configuration. The IRMPD spectrum of protonated Pro_3 matches nicely with the calculated spectrum of the lowest energy conformation identified here, having both amide bonds in the *cis* configuration. For Pro_4H^+, Pro_5H^+ and Pro_6H^+, the experimental spectra match with calculated structures having mixed *cis* and *trans* peptide bonds, however, the first amide bond was found to be *cis* in all cases. In these conformations, the protonated N-terminus can hydrogen-bond with the carbonyl oxygens; recall that apart from the terminal ends of the peptide, there exist no H-bond donors in polyproline peptides. Hence, establishing that the first peptide bond is generally in a *cis*-configuration supports the proposition of a diketopiperazine-b_2 fragment for protonated polyprolines.

Acknowledgements

The authors gratefully acknowledge Mr. S. Popovic and Dr. J. van Maarseveen from the University of Amsterdam for synthesizing the cyclo-ProPro model compound as well as Dr. D. Löwik from the Radboud University for synthesizing the Pro_2 and Pro_6 peptides. They gratefully acknowledge the FELIX staff, particularly Dr. A.F.G. van der Meer, Dr. B. Redlich and Mr. R. van Buuren for technical support. Financial support for this project was provided by NWO Chemical Sciences under VICI project nr. 724.011.002. The authors also thank NWO Physical Sciences (EW) and the SARA Supercomputer Center for providing the computational resources. This work is part of the research program of FOM, which is financially supported by NWO.

REFERENCES

(1) Huang, Y.; Tseng, G. C.; Yuan, S.; Pasa-Tolic, L.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *A Data-Mining Scheme for Identifying Peptide Structural Motifs Responsible for Different MS/MS Fragmentation Intensity Patterns* J. Proteome Res. **2007**, *7*, 70.

(2) Li, W.; Ji, L.; Goya, J.; Tan, G.; Wysocki, V. H. *SQID: An Intensity-Incorporated Protein Identification Algorithm for Tandem Mass Spectrometry* J. Proteome Res. **2011**, *10*, 1593.

(3) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *Mobile and localized protons: A framework for understanding peptide dissociation.* J. Mass Spectrom. **2000**, *35*, 1399.

(4) Paizs, B.; Suhai, S. *Fragmentation pathways of protonated peptides* Mass Spectrom. Rev. **2005**, *24*, 508.

(5) Cox, K. A.; Gaskell, S. J.; Morris, M.; Whiting, A. *Role of the site of protonation in the low-energy decompositions of gas-phase peptide ions* J. Am. Soc. Mass Spectrom. **1996**, *7*, 522.

(6) Boyd, R.; Somogyi, Á. *The Mobile Proton Hypothesis in Fragmentation of Protonated Peptides: A Perspective* J. Am. Soc. Mass Spectrom. **2010**, *21*, 1275.

(7) Dongré, A. R.; Jones, J. L.; Somogyi, Á.; Wysocki, V. H. *Influence of Peptide Composition, Gas-Phase Basicity, and Chemical Modification on Fragmentation Efficiency: Evidence for the Mobile Proton Model* J. Am. Chem. Soc. **1996**, *118*, 8365.

(8) Paizs, B.; Suhai, S. *Towards Understanding the Tandem Mass Spectra of Protonated Oligopeptides. 1: Mechanism of Amide Bond Cleavage* J. Am. Soc. Mass Spectrom. **2004**, *15*, 102.

(9) Yalcin, T.; Khouw, C.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. *Why are b ions stable species in peptide spectra?* J. Am. Soc. Mass Spectrom. **1995**, *6*, 1164.

(10) Yalcin, T.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. *The structure and fragmentation of b_n ($n \geq 3$) ions in peptide spectra* J. Am. Soc. Mass Spectrom. **1996**, *7*, 233.

(11) Harrison, A. G.; Csizmadia, I. G.; Tang, T.-H. *Structure and fragmentation of b_2 ions in peptide mass spectra* J. Am. Soc. Mass Spectrom. **2000**, *11*, 427.

(12) Paizs, B.; Suhai, S. *Combined quantum chemical and RRKM modeling of the main fragmentation pathways of protonated GGG. I. Cis-trans isomerization around protonated amide bonds* Rapid Commun. Mass Spectrom. **2001**, *15*, 2307.

(13) Paizs, B.; Suhai, S. *Combined quantum chemical and RRKM modeling of the main fragmentation pathways of protonated GGG. II. Formation of b_2, y_1, and y_2 ions* Rapid Commun. Mass Spectrom. **2002**, *16*, 375.

(14) Gucinski, A. C.; Chamot-Rooke, J.; Nicol, E.; Somogyi, Á.; Wysocki, V. H. *Structural Influences on Preferential Oxazolone versus Diketopiperazine b_2^+ Ion Formation for Histidine Analogue-Containing Peptides* J. Phys. Chem. A **2012**, *116*, 4296.

(15) Perkins, B. R.; Chamot-Rooke, J.; Yoon, S. H.; Gucinski, A. C.; Somogyi, Á.; Wysocki, V. H. *Evidence of Diketopiperazine and Oxazolone Structures for HA b_2^+ Ion* J. Am. Chem. Soc. **2009**, *131*, 17528.

(16) Pal, D.; Chakrabarti, P. *Cis peptide bonds in proteins: residues involved, their conformations, interactions and locations* J. Mol. Biol. **1999**, *294*, 271.

(17) Jabs, A.; Weiss, M. S.; Hilgenfeld, R. *Non-proline cis peptide bonds in proteins* J. Mol. Biol. **1999**, *286*, 291.

(18) Nold, M. J.; Cerda, B. A.; Wesdemiotis, C. *Proton affinities of the N- and C-terminal segments arising upon the dissociation of the amide bond in protonated peptides* J. Am. Soc. Mass Spectrom. **1999**, *10*, 1.

(19) Brandts, J. F.; Brennan, M.; Lin, L.-N. *Unfolding and refolding occur much faster for a proline-free proteins than for most proline-containing proteins* Proc. Nat. Acad. Sci. **1977**, *74*, 4178.

(20) Bleiholder, C.; Suhai, S.; Harrison, A.; Paizs, B. *Towards Understanding the Tandem Mass Spectra of Protonated Oligopeptides. 2: The Proline Effect in Collision-Induced Dissociation of Protonated Ala-Ala-Xxx-Pro-Ala (Xxx = Ala, Ser, Leu, Val, Phe, and Trp)* J. Am. Soc. Mass Spectrom. **2011**, *22*, 1032.

(21) Breci, L. A.; Tabb, D. L.; Yates, J. R.; Wysocki, V. H. *Cleavage N-Terminal to Proline: Analysis of a Database of Peptide Tandem Mass Spectra* Anal. Chem. **2003**, *75*, 1963.

(22) Grewal, R. N.; El Aribi, H.; Harrison, A. G.; Siu, K. W. M.; Hopkinson, A. C. *Fragmentation of Protonated Tripeptides: The Proline Effect Revisited* J. Phys. Chem. B **2004**, *108*, 4899.

(23) Harrison, A. G.; Young, A. B. *Fragmentation reactions of deprotonated peptides containing proline. The proline effect* J. Mass Spectrom. **2005**, *40*, 1173.

(24) Schwartz, B. L.; Bursey, M. M. *Some proline substituent effects in the tandem mass spectrum of protonated pentaalanine* Biol. Mass Spectrom. **1992**, *21*, 92.

(25) Vaisar, T.; Urban, J. *Probing Proline Effect in CID of Protonated Peptides* J. Mass Spectrom. **1996**, *31*, 1185.

(26) Loo, J. A.; Edmonds, C. G.; Smith, R. D. *Tandem mass spectrometry of very large molecules. 2. Dissociation of multiply charged proline-containing proteins from electrospray ionization* Anal. Chem. **1993**, *65*, 425.

(27) Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R. *Protein sequencing by tandem mass spectrometry* Proc. Nat. Acad. Sci. **1986**, *83*, 6233.

(28) Martin, S. A.; Biemann, K. *A comparison of keV atom bombardment mass spectra of peptides obtained with a two-sector mass spectrometer with those from a four-sector tandem mass spectrometer* Int. J. Mass Spectrom. and Ion Proc. **1987**, *78*, 213.

(29) Tang, X. J.; Thibault, P.; Boyd, R. K. *Fragmentation reactions of multiply-protonated peptides and implications for sequencing by tandem mass spectrometry with low-energy collision-induced dissociation* Anal. Chem. **1993**, *65*, 2824.

(30) Vaisar, T.; Urban, J. *Gas-phase fragmentation of protonated mono-N-methylated peptides. Analogy with solution-phase acid-catalyzed hydrolysis* J. Mass Spectrom. **1998**, *33*, 505.

(31) Kapp, E. A.; Schütz, F.; Reid, G. E.; Eddes, J. S.; Moritz, R. L.; O'Hair, R. A. J.; Speed, T. P.; Simpson, R. J. *Mining a Tandem Mass Spectrometry Database To Determine the Trends and Global Factors Influencing Peptide Fragmentation* Anal. Chem. **2003**, *75*, 6251.

(32) Unnithan, A. G.; Myer, M. J.; Veale, C. J.; Danell, A. S. *MS/MS of Protonated Polyproline Peptides: The Influence of N-terminal Protonation on Dissociation* J. Am. Soc. Mass Spectrom. **2007**, *18*, 2198.

(33) Moradi, M.; Babin, V.; Roland, C.; Darden, T. A.; Sagui, C. *Conformations and free energy landscapes of polyproline peptides* Proc. Nat. Acad. Sci. **2009**, *106*, 20746.

(34) Counterman, A. E.; Clemmer, D. E. *Anhydrous Polyproline Helices and Globules* J. Phys. Chem. B **2004**, *108*, 4885.

(35) Oomens, J.; Young, S.; Molesworth, S.; Stipdonk, M. *Spectroscopic evidence for an oxazolone structure of the b_2 fragment ion from protonated tri-alanine* J. Am. Soc. Mass Spectrom. **2009**, *20*, 334.

(36) Durand, S.; Rossa, M.; Hernandez, O.; Paizs, B.; Maître, P. *IR Spectroscopy of b_4 Fragment Ions of Protonated Pentapeptides in the X–H (X = C, N, O) Region* J. Phys. Chem. A **2013**, *117*, 2508.

(37) Yoon, S. H.; Chamot-Rooke, J.; Perkins, B. R.; Hilderbrand, A. E.; Poutsma, J. C.; Wysocki, V. H. *IRMPD spectroscopy shows that AGG forms an oxazolone b_2^+ ion* J. Am. Chem. Soc. **2008**, *130*, 17644.

(38) Grzetic, J.; Oomens, J. *Spectroscopic Identification of Cyclic Imide b_2-Ions from Peptides Containing Gln and Asn Residues* J. Am. Soc. Mass Spectrom **2013**, *24*, 1228.

(39) Valle, J. J.; Eyler, J. R.; Oomens, J.; Moore, D. T.; van der Meer, A. F. G.; von Helden, G.; Meijer, G.; Hendrickson, C. L.; Marshall, A. G.; Blakney, G. T. *Free electron laser-Fourier transform ion cyclotron resonance mass spectrometry facility for obtaining infrared multiphoton dissociation spectra of gaseous ions* Review of Scientific Instruments **2005**, *76*.

(40) Friedrich, A.; Jainta, M.; Nieger, M.; Bräse, S. *One-Pot Synthesis of Symmetrical and Unsymmetrical Diketopiperazines from Unprotected Amino Acids* Synlett **2007**, *2007*, 2127.

(41) D.A. Case, T. A. D., T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Goetz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman *AMBER 2012* **2012**.

(42) Stedwell, C. N.; Galindo, J. F.; Gulyuz, K.; Roitberg, A. E.; Polfer, N. C. *Crown Complexation of Protonated Amino Acids: Influence on IRMPD Spectra* J. Phys. Chem. A **2012**, *117*, 1181.

(43) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Wallingford CT, 2009.

(44) Sediki, A.; Snoek, L. C.; Gaigeot, M.-P. *N–H+ vibrational anharmonicities directly revealed from DFT-based molecular dynamics simulations on the* Ala_7H^+ *protonated peptide* Int. J. Mass Spectrom. **2011**, *308*, 281.

(45) Cimas, A.; Vaden, T. D.; de Boer, T. S. J. A.; Snoek, L. C.; Gaigeot, M. P. *Vibrational Spectra of Small Protonated Peptides from Finite Temperature MD Simulations and IRMPD Spectroscopy* J. Chem. Theory Comput. **2009**, *5*, 1068.

(46) Leavitt, C. M.; DeBlase, A. F.; Johnson, C. J.; van Stipdonk, M.; McCoy, A. B.; Johnson, M. A. *Hiding in Plain Sight: Unmasking the Diffuse Spectral Signatures of the Protonated N-Terminus in Isolated Dipeptides Cooled in a Cryogenic Ion Trap* J. Phys. Chem. Lett. **2013**, *4*, 3450.

Chapter 9.

Observation of anionic diketopiperazine b_2 fragment ions: an ion spectroscopy study

In the previous chapter we described a set of results detailing the conformations of a series of protonated polyproline peptides. It was shown that the b_2 fragment of Pro_3H^+ is characteristic of the thermodynamically-favoured diketopiperazine structure, showing no absorption bands higher than 1800 cm^{-1} where an oxazolone carbonyl stretch would typically be found. As well, a series of calculated structures for Pro_3H^+, Pro_4H^+, Pro_5H^+ and Pro_6H^+ were identified having *cis* configuration of their first peptide bond in all cases, which is mechanistically consistent with the formation of diketopiperazine-b_2 fragments for protonated polyprolines.

Here, we analyze a similar set of proline-based peptides as anions. While charge-directed fragmentation for protonated peptides depends on the position and dynamics of the extra proton, the analogous fragmentation mechanisms for peptide anions is based on the site of deprotonation. After activation of peptide anions, are protons mobile in a similar fashion as in protonated peptides? Polyprolines offer a unique insight into this question as they possess no amide protons nor any acidic side chain protons; the C-terminal proton is the only acidic proton and the N-terminal proton is the only mobilizable proton.

Here, we observe, for the first time, fragmentation of anionic peptides along the diketopiperazine pathway, as well as a unique fragmentation pattern of Pro_{2-7}^- upon collisional activation. Protonated histidine containing peptides have also been noted to produce diketopiperazine b_2 fragement and we will explore the behavior of their deprotonated analogs.*

---

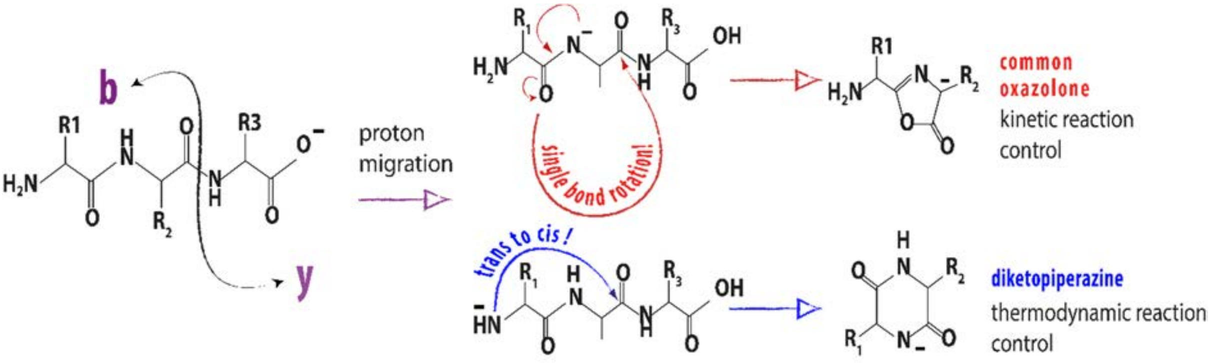
* J. Grzetic, J.K. Martens, and J. Oomens, *Observation of anionic diketopiperazine b_2 fragment ions – an ion spectroscopy study*, manuscript in preparation

9.1. Introduction

The field of modern proteomics largely centres on mass spectrometric methods[1-3]. This is mainly due to their unmatched sensitivity, the ability to isolate and analyze single analytes from what are often very complex biological matrices, and the ability to examine compounds of masses spanning many orders of magnitude (<100 – >10^6 Da). As well, advances in modern mass spectrometers have allowed protein analysis to become relatively routine and inexpensive. As such, sequencing and identification of proteins using fragment mass patterns and known databases (established from DNA sequences) allows near automation of the process in many cases[4]. Unfortunately, less than certain identification frequently results from incomplete sequence coverage being obtained from the fragmentation information. This is partly due to the incomplete understanding of the fragmentation chemistry involved in peptide dissociation[5]. Were a more complete description of peptide dissociation patterns incorporated in sequence analysis based on mass fragmentation patterns, better sequence coverage and more certain identification would undoubtedly result. This includes the incorporation of non-standard fragmentation patterns as well as relative intensities within a given fragmentation pathway and between different pathways[6,7]. Based on chemical intuition and a large amount of data from tandem MS experiments, various peptide dissociation pathways and fragment structures have been suggested in the literature. However, additional computational and experimental evidence is typically required to firmly establish the suggested reaction pathways and fragment structures, especially since various isomeric and tautomeric structures are often conceivable. In addition to information obtained from computational results, experimental evidence, especially using ion spectroscopy and ion mobility, has recently been successful for determining fragment structures. Combining experimental IR spectra with results of computational chemistry (or with experimental spectra of model compounds) the structure of peptide fragment ions can often be identified with a high degree of confidence[8,9].

Until now, spectroscopic studies of collision induced dissociation (CID) products from protonated peptides have largely focused on *a*- and *b*-type fragments[10-12]. A notable success of the technique can be seen in the sequence dependent nature of b_2-type fragments, having either oxazolone or diketopiperazine structures[13,14]. Residue specific fragmentation of peptides containing glutamine or asparagine residues has also been demonstrated recently. In contrast to the more common formation of oxazolone (and rarely diketopiperazine) b_2-type structures, the presence of a Asn or Gln residue, when at the second position from the N-terminus, was found to generate b-type ions via nucleophilic attack of the side chain amide nitrogen atom of Asn and Gln on the adjacent peptide carbonyl carbon[15]. This was observed to lead to formation of succinimide and glutarimide ring structures for these b_2-type ions, respectively. In addition to the fragment structures above, ion spectroscopy was also instrumental in the proposition that larger b-ions may undergo "head-to-tail" macrocyclization, later reopening and generating fragments with a scrambled sequence[16,17].

Fragmentation of protonated peptides most commonly follows the proposed mobile proton model, where ion activation mobilizes the thermodynamically trapped proton to relocalize at sites of lower basicity[18]. If this less basic site is a backbone amide group, the amide bond weakens and increases the susceptibility of the carbonyl carbon to nucleophilic attack. Post-nucleophilic attack cleavage of the weakened C(=O)-NH_2^+ bond corresponds to the common b_n-y_m pathway. To form the commonly observed oxazolone b-fragments, the nucleophile in this reaction is one of the backbone carbonyl groups. For formation of the uncommon diketopiperazine b-type fragment, the nucleophile is the nitrogen of the N-terminus. Considering the sterics of this reaction, one can see that the oxazolone generating reaction requires a *trans* conformation of the first peptide bond, while the corresponding reaction to form the diketopiperazine requires the *cis* conformation (Scheme 9.1). As most peptides bonds are most stable in the *trans* conformation, the high barrier to isomerize to the *cis* conformation is thought to be the reason that the oxazolone-type structure is more common (kinetic control)[19-21]. Only in specific cases is isomerization from *trans* to *cis* conformation of the peptide bond thought to be feasible, relating to the partially reduced bond order of the peptide bond after protonation of the amide nitrogen[22].



Scheme 9.1. Possible dissociation/rearrangement reactions based on the 'mobile proton' model, leading to different b_2-ion structures, for deprotonated peptides.

Proline-containing peptides represent a unique case, as more frequent *cis* configuration of amide bonds involving proline residues have been demonstrated. The observation of diketopiperazine b_2 fragments from such peptides has as well been demonstrated using ion spectroscopy. In solution, amide bonds not involving proline are thought to have *cis* peptide bonds in only 0.03-0.05 % of cases due to both the difference in relative energies of the *cis/trans* conformers (*cis* ~+10 kJ mol^{-1}) and the barrier to isomerization (~85 kJ mol^{-1}). For amide bonds containing a proline residue, however, the *cis/trans* conformers are only separated by ~2 kJ mol^{-1} and the isomerization has a barrier of approximately ~55 kJ mol-1[23,24]. The unique conformations of peptides containing proline residues are related to its pyrrolidine ring structure and the particular rigidity it brings to the peptide backbone. In addition, enhanced fragmentation at Xxx-Pro amide bonds has been frequently observed, commonly referred to as the proline effect[25-28]. (Ref)

While MS/MS-based peptide sequencing has almost exclusively focused on the fragmentation of protonated peptides (with some exceptions) fragmentation of deprotonated peptides presents a few promising and complementary features in comparison to conventional experiments[29-31]. Firstly, anionic peptides typically generate different fragmentation patterns and thus provide additional sequence information. As well, molecular ions from peptides containing many acidic residues or phosphorylation, can be more efficiently generated by deprotonation than by protonation. In contrast to the mainly a-, b- and y-type backbone fragments generated from protonated peptides, peptide anions additionally exhibit c- and z-type ions in many cases. Various studies have therefore addressed the mechanisms of fragmentation of deprotonated peptides using tandem MS and computational methods[30-32].

The first demonstration of IR spectroscopic characterization of deprotonated peptide fragments considered the a_3 fragment from Ala_3, focusing on identifying the deprotonation site of the N-terminal fragment no longer having a carboxylate group (C-terminus)[33]. Assuming that fragmentation of deprotonated peptides is as well normally charge-directed, this is an important issue to clarify. These results showed that the a_3^- fragment is an amidate, as previously postulated.

As observed for protonated peptides, amide bond cleavage is a common fragmentation channel for deprotonated peptides, leading to the formation of b- or y-type ions[34-36]. What is less clear is whether the structures of these b-type fragments from deprotonated peptides correspond to those generated from protonated peptides. For singly protonated peptides, the mass-spectrometrically observed fragment on this pathway is the one retaining the proton after the two products dissociate (i.e. influenced by relative gas phase basicities). For deprotonated peptides, the corresponding considerations would have to focus on which of the fragments has a higher gas-phase acidity, or can better stabilize the negative charge. Do they also form mainly oxazolone and rarely diketopiperazine structures as is now well established for protonated peptides? It is moreover not entirely clear where the N-terminal fragments deprotonate in the absence of acidic residues. Recently, the b_2 fragment from deprotonated Ala_3 and deprotonated AlaTyrAla have been spectroscopically shown to be oxazolones [36], as had been previously suggested [9, 36]. c-type ions are also commonly observed in CID mass spectra of anionic peptides containing alkyl [37], acidic[32,38,39], hydroxyl[40], basic[31] and aromatic[41] residues. These c-fragments clearly provide information that is not available from the corresponding positive-ion experiment, allowing additional sequencing from the N-terminus.

Here, we will address some of the questions surrounding the conformations and fragmentation behavior of small deprotonated peptides containing proline and histidine using a combined experimental/computational approach. More specifically we use an MM/MD/DFT procedure to match calculated peptide and fragment IR spectra and structures with those experimentally obtained using ion spectroscopy. We have considered the structures of Pro_3^- and its CID generated b_2^- and a_2^- fragments as well as $[AlaHisAla - H]^-$

and its b_2^- fragment. These results constitute the first demonstration of an anionic diketopiperazine b_2^- peptide fragment.

9.2. Experimental and Computational Methods

9.2.1. Infrared Multiple Photon Dissociation (IRMPD) Spectroscopy

Infrared spectra of deprotonated peptide and peptide fragment ions were obtained using a previously described Fourier transform-ion cyclotron resonance (FT-ICR) mass spectrometer and the tunable infrared radiation of the FELIX free electron laser[42]. Deprotonated molecular ions ($[M-H]^-$) ions were electrosprayed from millimolar solutions in 50:50 acetonitrile:water containing ~0.1% NH_4OH. Peptides were obtained from GeneCust (Luxemburg) and Bachem (Bubendorf, Switzerland) and used as received. Pro_2, Pro_6 and Pro_7 were synthesized at Radboud University Nijmegen by Dennis Löwik. Once electrosprayed, $[M-H]^-$ ions were accumulated in a linear hexapole trap before being directed to the ICR cell. Peptide fragment ions were either generated by nozzle-skimmer dissociation in the high-pressure region preceding the hexapole trap, or in-cell by photodissociation using a 35 W continuous-wave CO_2 laser for some hundreds of milliseconds (ULR-25; Universal Laser Systems, Scottsdale, USA).

Once trapped in the cell, isolation of the ion of interest is accomplished using a stored waveform inverse Fourier Transform (SWIFT) excitation pulse. Isolated ions were then irradiated by the tunable infrared radiation from FELIX, arriving as 5 µs macropulses (5 Hz) of approximately 30-40 mJ and having a bandwidth of ~0.4% of the central frequency. When the IR energy is resonant with one of the vibrational modes of the system, absorption leads to an increase in the internal energy and results in unimolecular dissociation. The result is a frequency-dependent fragmentation signal. Relation of these parent and fragment ion intensities as the fragmentation yield ($\Sigma I(\text{fragment ions})/\Sigma I(\text{parent + fragment ions})$) generates an infrared vibrational action spectrum. All spectra are linearly corrected for laser power and the frequency is calibrated using a grating spectrometer.

Collision induced dissociation (CID) experiments were obtained from a Bruker AmaZon Speed ETD quadrupole ion trap. $[M-H]^-$ ions were generated using electrospray ionization (ESI) from approximately 10 micromolar peptide solutions in 50:50 acetonitrile:water with 0.1% NH_4OH. After isolation in the trap, the $[M-H]^-$ ions were collisionally activated for 40 ms using an amplitude parameter of ~0.5-1.0 V.

9.2.2. Computational procedure

An MM/MD/DFT computational procedure was used analogously to that described in Chapter 8 and only a brief overview will be given here. The molecular dynamics approach is necessary because, even for small peptides, the conformational landscapes are complex. As a result, manual definition of input geometries using chemical intuition is often insufficient. We have therefore employed a molecular mechanics/molecular dynamics (MM/MD) approach using AMBER 12[43]. Initial guess geometries were first optimized in the Gaussian09 package[44] at the B3LYP/6-31++G(d,p) level of theory. Charges from these initial results were used for parameterization of the nonstandard terminations of the peptides. A simulated annealing procedure (up to 1000 K) was then applied. Within this temperature range it was found that *cis/trans* isomerization of the peptide bonds was accessible. The resulting conformations were grouped based on their structures and lead to approximately 30 structures for each peptide or fragment considered. These structures were each optimized at the B3LYP/6-31++G(d,p) level followed by vibrational analysis within the harmonic oscillator model (vibrations were scaled by 0.975). Calculated line spectra were broadened using a Gaussian function with a FWHM of 25 cm^{-1} to facilitate comparison with experiment.

9.3. Results and discussion

9.3.1. Structure of the deprotonated Pro_3 precursor peptide

Comparison of the experimental IRMPD spectrum of deprotonated Pro_3 with the calculated spectra of structures deemed to be reasonable candidate structures are presented in Figures 9.1 and 9.2. Previously it was shown, in the work of Oomens and coworkers, that most amino acids and short peptides, when singly deprotonated, preferentially deprotonate at the C-terminus and form a carboxylate anion[33,36,45]. However, noting the conformational uniqueness of proline-based peptides and their lack of amide hydrogens, various alternative configurations and conformations were considered, including possible deprotonation of an α-carbon (leading to an enolate structure) and deprotonation at the N-terminus.

Upon IRMPD, deprotonated Pro_3 at m/z 308, primarily dissociates into two channels, 114 m/z and 165 m/z, corresponding to the y_1^- fragment and loss of C_2H_4 (y_1 being the major channel). These mass channels were used to calculate the yield with results summarised in Figure 9.1. From the graphs presented, it is clear that only the C-terminally deprotonated structure matches with the experiment. An MD-based conformational exploration for this particular deprotonated structure was conducted to verify that all possible conformations

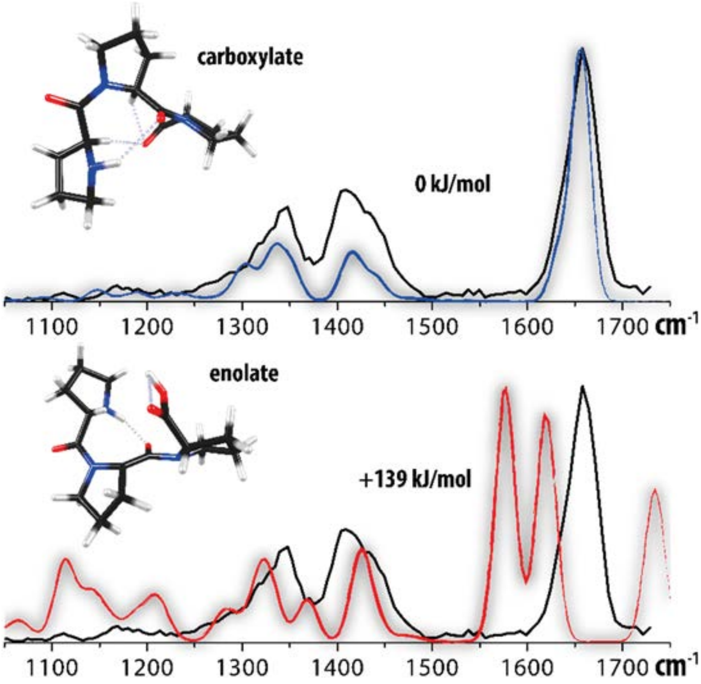


Figure 9.1. Comparison of the IRMPD spectrum of the deprotonated precursor peptide Pro_3^- (black trace) with DFT-calculated spectra (upper panel, blue trace) for structures deprotonated at the C-terminal carboxylate moiety and at the C_α (lower panel, red trace). Calculated relative zero point energies are given for each structure.

were considered, including conformations with both *cis* and *trans* peptide linkages. From a detailed inspection of the six comparisons in Figure 9.2, one observes that all calculated spectra are relatively similar. Nevertheless, it appears that the calculated spectrum of conformer c0 matches most closely with experiment; all vibrational modes are reproduced nearly exactly. As well, c0 is the lowest energy structure, and it is worth noting that within this conformer both peptide bonds are *cis* configured (consistent with its protonated counterpart, see Chapter 8). What is noticeable in the other five panels containing spectra for conformers c2, c3, c1, c11 and c8, is that none of them reproduce the carbonyl stretching mode (a very characteristic stretching vibration) as closely as c0 does. Additionally, the lowest energy conformer having all *trans* peptide bonds is 9 kJ/mol higher in energy. Nevertheless, a small contribution to the ion population from other conformers (particularly c_2 and c_3) remains possible. From these observations, we now have a basis for considering the fragmentation pathways and structures formed from this anionic peptide.

9.3.2. The structure of the b_2 anion from deprotonated Pro_3

The b_2 anion was generated by CID from deprotonated Pro_3. As mentioned in previous chapters, the main structural candidates for this type of fragment are diketopiperazines and oxazolones. A comparison of theoretical and experimental IR spectra is presented in Figure 9.3, including the lowest energy structures for each family.

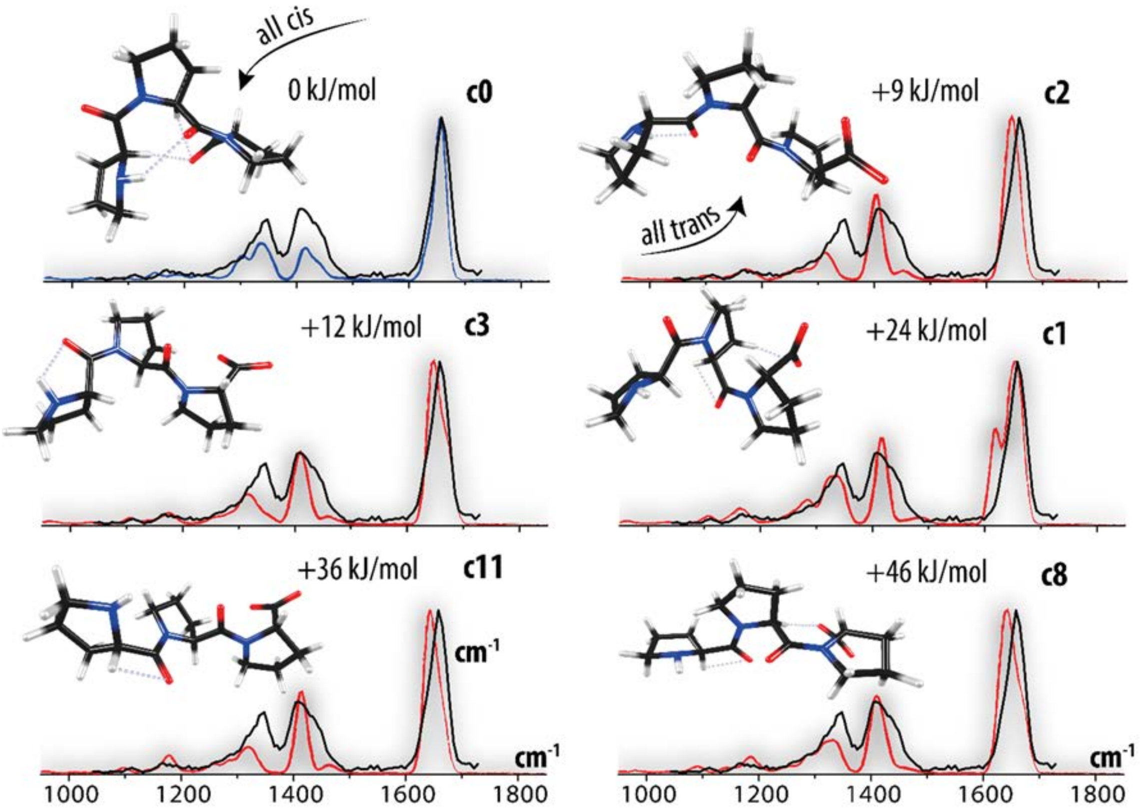


Figure 9.2. Calculated structures and spectra for $[Pro_3 - H]^-$. Clearly, structure c0 provides the best match with the experiment and it is the lowest energy conformer. Furthermore, we can assign all peptides bonds to be *cis*-configured in contrast to the structure c2 for which the match is somewhat worse, but all peptides bonds are *trans*.

From an analysis of the spectral comparisons in Figure 9.3 we conclude that only the calculated spectrum for the diketopiperazine matches the experiment. This is analogous with the b_2-type fragment from protonated Pro_3, also clearly found to be a diketopiperazine (Chapter 8). The structure in the middle panel of Figure 9.3 is referred to as a spiro – cyclopentanedione and can result from nucleophilic attack by the first C_α when deprotonated. Not only does this intuitively seem unlikely, its computed spectrum fails to match with the experimental one. The bottom panel of Figure 9.3 shows a comparison with the theoretical spectrum of the lowest energy oxazolone structure. It also clearly fails to match the experiment and is much higher in energy than both of the other structures.

Possible mechanistic pathways are presented in Scheme 9.2. In consideration of possible fragmentation pathways, the initial structure of the peptide has a significant influence. For $[Pro_3 - H]^-$, we have assigned the C-terminus as the site of deprotonation, leaving the proton of the N-terminus as the only acidic proton in the peptide. Virtually all other peptides contain protons that are relatively acidic, such as on specific residues or N-terminal or backbone amide protons, which can to some extent stabilize the negative charge in the gas phase. Once activated, this allows for proton migration to the C-terminus leading to an increased nucleophilicity elsewhere in the peptide, facilitating various reactions leading to

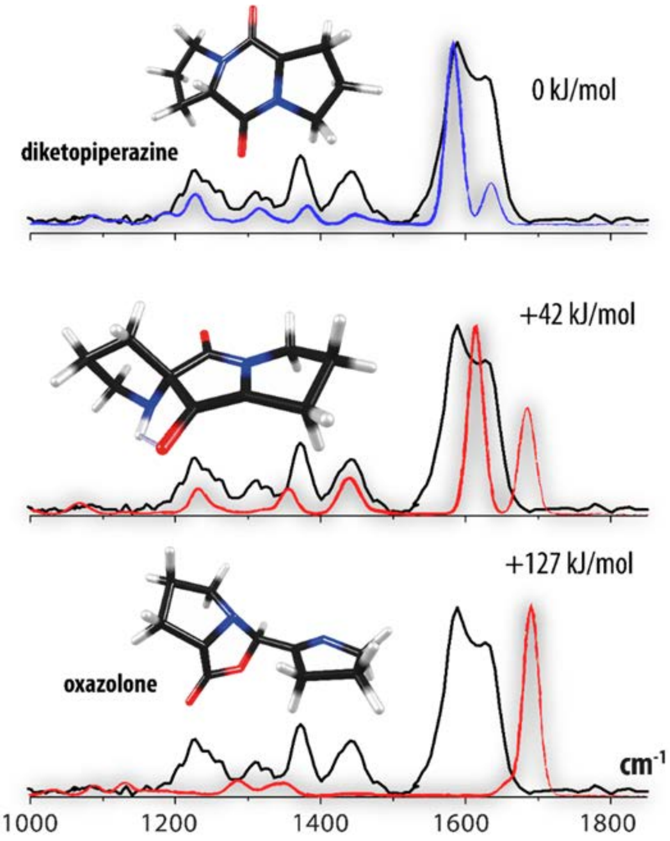
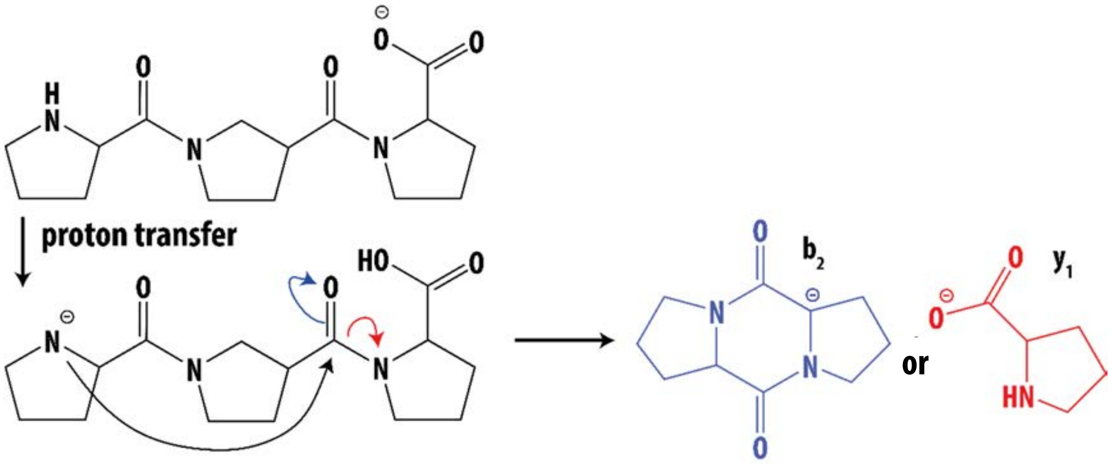


Figure 9.3. The IRMPD spectrum of the b_2-ion from Pro_3^- (black in all panels) compared to DFT-calculated spectra (blue and red) for isomeric candidate structures: diketopiperazine deprotonated on the α-carbon (upper panel), the spiro – cyclopentanedione derivative deprotonated at the C_α (middle panel) and oxazolone deprotonated on the C_α carbon. Calculated relative ZPE-corrected energies are given for each of the structures.

fragmentation along the different pathways. Such a scenario suggests that the 'mobile proton model' is applicable to deprotonated peptides as well. In the work of Bache and coworkers[46] scrambling of labelled protons in various anionic peptides was observed. However, for the polyproline peptides under study here, the possible deprotonation sites are very limited as there are no acidic side chain or backbone amide protons. As we have already assigned the structure of $[Pro_3 - H]^-$ to be deprotonated at the C-terminus, the only reasonable intramolecular proton-transfer reaction is with the N-terminus.

As presented in Chapter 5, formation of the b_2 ion from deprotonated Ala_3 and AlaTyrAla proceeds via deprotonation of a peptide bond NH, making the adjacent carbonyl oxygen a stronger nucleophile. As all peptides bonds are in the *trans* configuration, formation of an oxazolone structure requires only single-bond rotation and is favoured by reaction kinetics. Activation of the carbonyl, in a similar manner, for deprotonated Pro_3 is not possible as there are no amide protons. Moreover, we found that all peptide bonds for deprotonated Pro_3 are *cis*, which allows for easy formation of a diketopiperazine structure. Thus, assuming proton transfer to occur from the N-terminus to the C-terminus upon collisional activation, the N-terminus becomes a strong nucleophile and diketopiperazine formation is plausible, as suggested in Scheme 9.2 (see further discussion in Section 9.3.4 on CID MS/MS results).



Scheme 9.2. Mechanism for the formation of b_2/y_1 fragments from deprotonated Pro_3. The first step is proton transfer from the N-terminal nitrogen to the C-terminal carboxylate. The second step is a nucleophilic attack of the deprotonated terminal nitrogen on the amide carbonyl carbon, generating the diketopiperazine b_2 or y_1 anion.

Upon IRMPD, the isolated b_2 fragment (193 m/z) dissociates mainly to m/z 124 and 165 (dominant). The former corresponds to the loss of 1-pyroline and while the latter may either correspond to loss of CO, forming the a_2-ion, or C_2H_4. Loss of C_2H_4 may proceed via the fragmentation pathway suggested in Scheme 9.3.

9.3.3. Spectroscopic identification of the m/z 165 fragment

Loss of C_2H_4 was previously suggested in the fragmentation chemistry of anionic proline containing peptides by Harrison and coworkers[28]. While they suggested this neutral loss to occur from the a_2 fragment of deprotonated ProAla, the pathway suggested in Scheme 9.3 is based on the same chemical bond rearrangements within the proline ring. We have

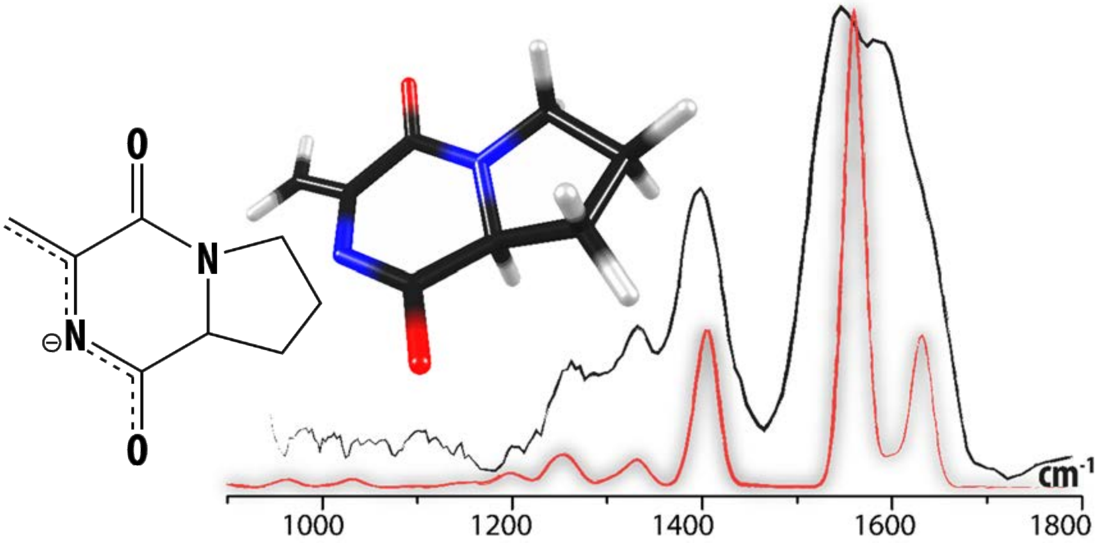
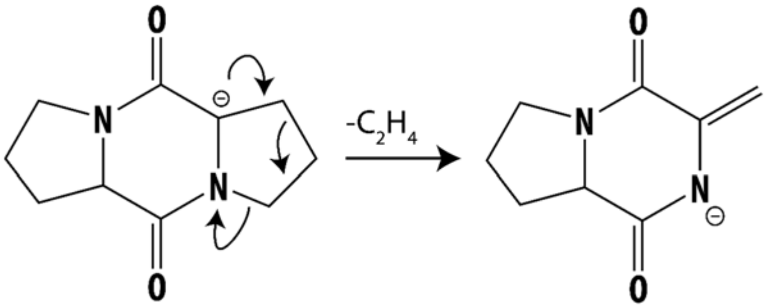


Figure 9.4. IRMPD spectrum of the 165 m/z anion from deprotonated Pro_3 (black) compared to the DFT-calculated spectrum (red). The negative charge is delocalized over five atoms within the 165 m/z ion, in contrast to three in the diketopiperazine, further supporting the proposition of this pathway, since charge delocalisation plays a significant role in stabilisation of gas-phase ions.



Scheme 9.3. Suggested mechanism for formation of the 165 m/z fragment from deprotonated b_2.

obtained additional confirmation for this suggested mechanism (Scheme 9.3) and the structure of the m/z 165 ion using IRMPD spectroscopy, and the results are represented in Figure 9.4.

Based on the favorable agreement between experiment and theory, we identify the m/z 165 anion as the structure presented in Figure 9.4, and confirm the suggested fragmentation pathway (Scheme 9.3), consistent with the assigned b_2-fragment structure. The fact that the negative charge is delocalized over five atoms within the 165 m/z ion, in contrast to three in the diketopiperazine, provides additional support for this pathway, as delocalisation plays a significant role in stabilisation of an ion in the gas phase.

9.3.4. CID of Pro_{2-7}^-

Figure 9.5 presents CID generated fragmentation mass spectra of deprotonated Pro_{2-7}. Firstly, it is interesting to note that all spectra are dominated by b- and y-type fragments, while c- and z-type fragments, which are often observed from CID of deprotonated peptides, are completely absent. This is understandable based on the additional linkage formed between the amide nitrogen and C_α by the proline side chain. In all cases, the b_2/y_{n-2} fragmentation channel is dominant, leading to the b_2^- base peak (193 m/z), except for Pro_3^-Interestingly, for Pro_3^-, the b_2/y_{n-2} channel is still dominant, however the ratio of abundances of the two fragments is reversed compared to the other polyprolines; the fragment retaining the charge, and thus observed in the mass spectrum, is mainly the y_1 fragment. In addition, with an increasing number of residues, other fragments (mainly larger b-type fragments) begin to appear in the spectra. This is especially noticeable for Pro_7^-, but also for Pro_6^- and Pro_5^-. The MS/MS spectrum of $[Pro_2 - H]^-$ features a dominant b_2^- fragment, although the y_1^- fragment is also of significant intensity. This fragmentation channel possibly involves attack by the deprotonated N-terminus on the C-terminal COOH group, resulting in loss of an OH^- which abstracts a proton from the nearby α-carbon and finally results in loss of neutral H_2O. This supports our suggestion that upon activation of the C-terminally deprotonated structure, proton transfer from the N-terminus to the C-terminus is required for fragmentation. While no calculations were completed for Pro_{4-7}^-,

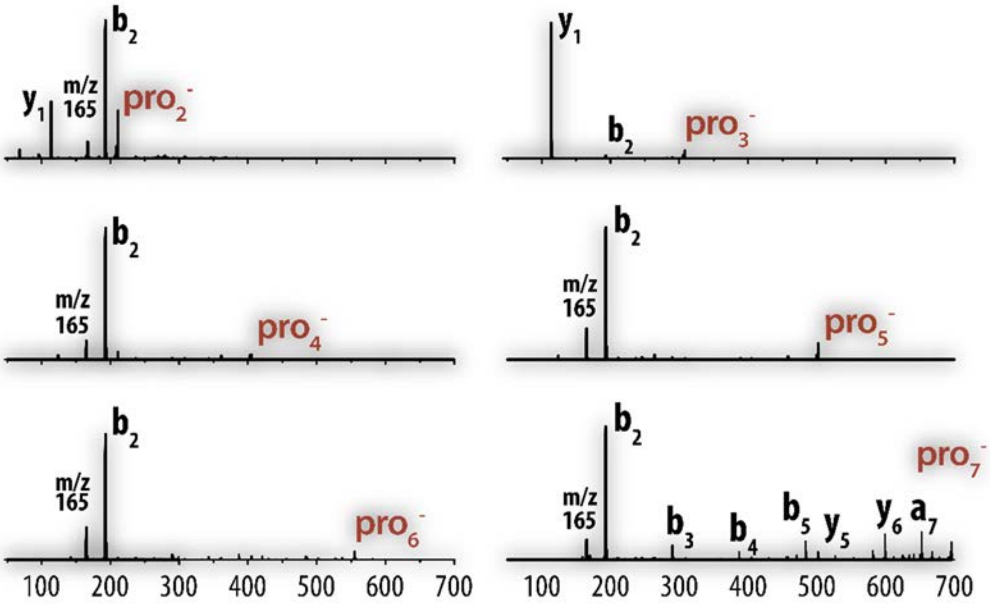


Figure 9.5 CID Mass spectra for $[Pro_n-H]^-$ (n=2-7) precursor ions. Labelled ions: y_6^- - 599 m/z , y_5^- - 501 m/z, y_1^- - 114 m/z, b_5^- - 484 m/z, b_4^- - 387 m/z, b_3^- - 290 m/z, b_2^- - 193 m/z, a_7^- - 650 m/z, 165 m/z (see text for description).

the fact that Pro_4^- and Pro_5^- produce only b_2^- fragments in the MS/MS spectra, but Pro_7^- (and to a smaller extent Pro_6^-) show additional b-fragments in their MS/MS spectra, suggests that the fragmentation mechanism is conformation-dependent. Both the structure for Pro_3^- that we have proposed, and those previously identified for $Pro_{3-6}H^+$ have *cis*-configured peptide bonds connecting the first and second residues. For polyproline anions, after proton transfer from the N-terminus to the C-terminus producing the reactive state of the system, the nucleophlic attack on the carbonyl of the second residue may then occur efficiently as it is close and accessible. Going to larger sizes, there is likely a distribution of conformers and proton transfer from N- to C-terminus is possibly entropically hindered, opening the possibility of additional fragmentation pathways involving other electrophiles.

9.3.5. Structure of deprotonated AlaHisAla precursor peptide

Comparison of the experimental IRMPD spectrum of deprotonated AHA with the calculated spectra of structures deemed reasonable candidates are presented in Figure 9.6. Similar to the analysis of deprotonated Pro_3, several sites of deprotonation in addition to the C-terminus were considered. However, the terminal COOH moiety was assigned as the deprotonation site also in this case. Even though deprotonation at the N-terminus was considered, in the geometry optimization the proton of the adjacent amide group (second) was abstracted, leading to the amidate structure shown in the bottom left panel of Figure 9.6. Not only was this structure calculated to be 80 kJ/mol higher in energy than the C-terminal carboxylate structure, its spectrum is in clear disagreement with the experiment and does not warrant further consideration. Deprotonation of the histidine ring was also considered (lower right panel in Fig. 9.6), but was again found to be high in energy and in disagreement with the experiment.

Four of the structures deprotonated on the C-terminus are presented in Figure 9.6. Of these, it is worth noting that structure c15 has a *cis* peptide linkage between the first and second residues. Structure c0 is a very similar conformation, except that the first peptide bond is now *trans*. While c15 is calculated to be +16 kJ/mol higher in energy than c0, we cannot spectrally distinguish these structures. c0 gives slightly better agreement in the amide I region, while c15 is slightly better in the amide II region (disregarding intensities). Structure c14, in which the N-terminus and C-terminus are hydrogen bound, is easily dismissed on

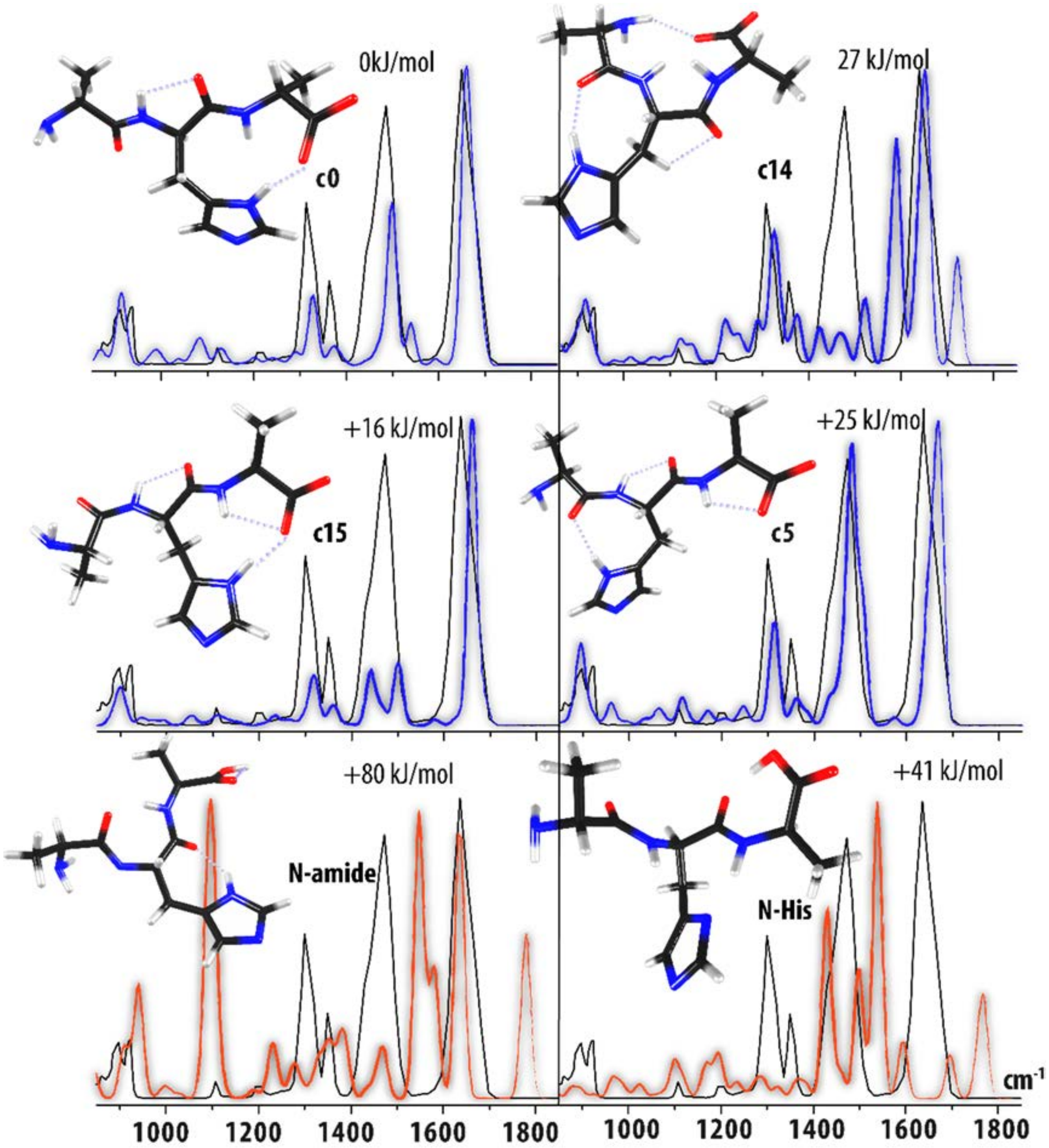


Figure 9.6. Comparison of the IRMPD spectrum of the deprotonated precursor peptide [AHA – H]$^-$ (black trace) with the DFT-calculated spectra (upper panel, blue trace) for c0, c14, c15 and c5 structures deprotonated at the C-terminal carboxylate moiety, amide N, and His side chain N (lower panel, red trace). Calculated relative zero point corrected energies are given for each structure.

both thermodynamic and spectral grounds. Structure c5 is similar to c0, except for the histidine ring being rotated by ~180^0 and the N-H of the ring being hydrogen bonded to the first C=O instead of to the C-terminus. While spectrally indistinguishable, c5 is calculated to be 25 kJ/mol higher in energy.

9.3.6. The structure of b_2 fragment of deprotonated AlaHisAla

We have additionally considered the structure of the b_2-fragment from [AHA – H]$^-$. Figure 9.7 compares various possible structures for this fragment ion, including oxazolone and diketopiperazine variants. It is immediately apparent that [AHA – H]$^-$ generates a diketopiperazine b_2^- fragment. Two low energy conformations of the oxazolone fragment are presented in Figure 9.7 and both are in clear disagreement with the experimental spectrum. Of the diketopiperazine fragments considered, one can see that only structures

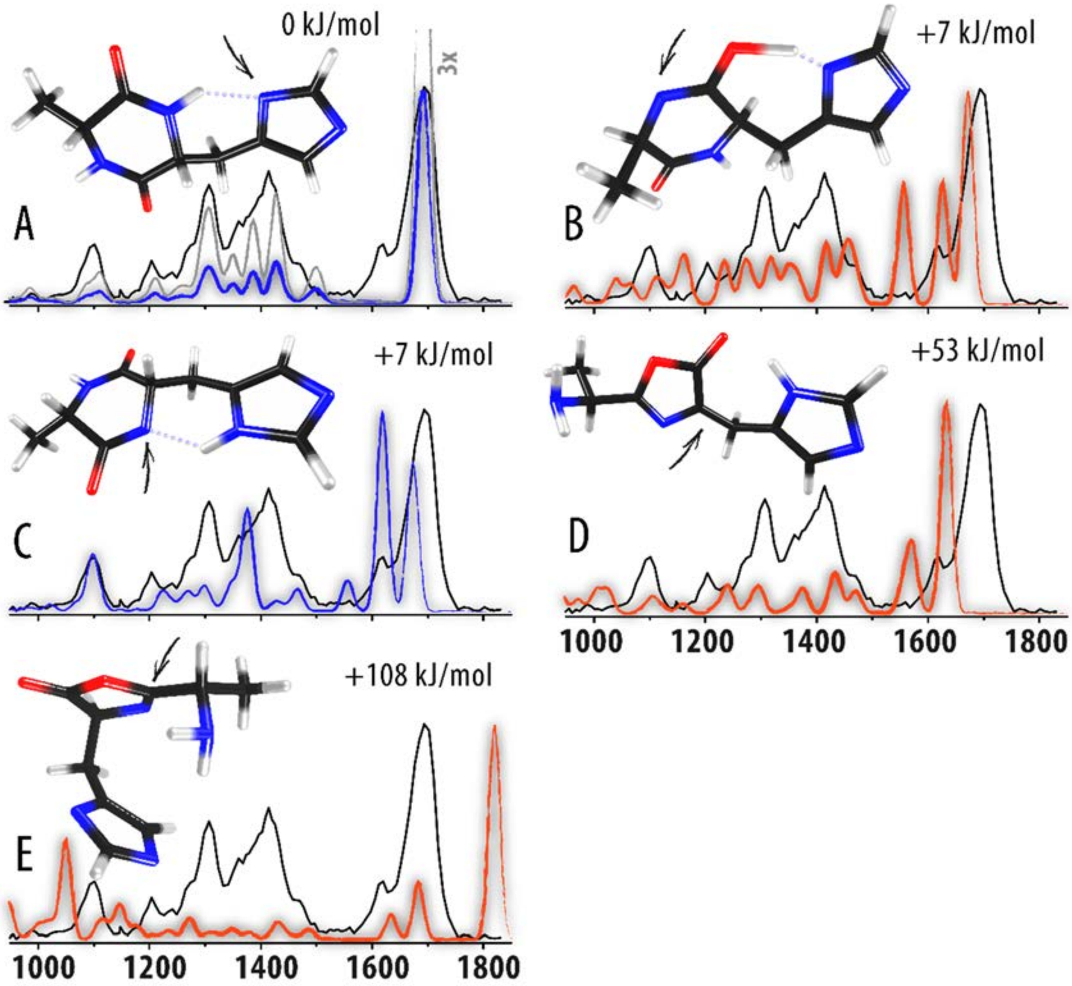


Figure 9.7. IRMPD spectrum of the b_2-ion from deprotonated AHA (black in all panels) compared to the DFT-calculated spectra (blue and red) for isomeric candidate structures: diketopiperazine deprotonated on His residue nitrogen atom (A), diketopiperazine deprotonated on the nitrogen adjacent to alanine residue (B), diketopiperazine deprotonated on nitrogen adjacent to His residue (C), oxazolone deprotonated on ring-embedded carbon adjacent to His residue (D) and oxazolone deprotonated on ring-embedded carbon adjacent to Ala residue. Calculated relative ZPE-corrected energies are given for each of the structures. The arrow points to the deprotonation site. For structure A, calculated spectral intensity is shown with 3x amplification to highlight the agreement.

A and C match with the experiment, but B is inconsistent. While, structure A is lowest energy, the slightly higher energy (+7 kJ/mol) structure C likely contributes the small second feature in the carbonyl stretching band at ~1625 cm^{-1}. As formation of a diketopiperazine b$_2$-fragment appears to be associated with *cis* conformation of the first peptide bond in the parent peptide, this observation provides additional insight into the the conformation of [AHA – H]$^-$ discussed above. Specifically, formation of a diketopiperazine b$_2$-fragment suggests that conformation c15 of AHA$^-$ is the dominant gas-phase conformation of AHA-, or that isomerization between c0 and c15 (i.e. rotation about the first amide bond, connecting the *cis* and *trans* conformations) is relatively easy, at least after activation.

9.4. Conclusions

To summarize, we have presented the first spectroscopic observation that anionic peptides can fragment along the diketopiperazine pathway. The b$_2$ fragment of Pro$_3^-$ was spectroscopically characterized and found to be consistent with a diketopiperazine-type structure, having no absorption bands higher than 1800 cm^{-1}, where one would expect the oxazolone carbonyl stretch. Noting this, we assign this ion to be one of the first cases where an anionic b-type fragment is not formed along the oxazolone pathway. We also observed that the CID MS/MS spectra of a series polyprolines (2-7 residues) contain no c- or z- type fragments, as is common in CID of anionic peptides. Instead, the CID spectra are dominated by the b$_2^-$ fragment ion. Only Pro$_3^-$ differed from this trend, as the y$_1^-$ fragment was essentially the only ion observed in the fragmentation spectrum. We have also characterized the b$_2^-$ from AHA$^-$, establishing its structure to be a diketopiperazine. Characterization of the parents Pro$_3^-$ and AHA$^-$ using IRMPD, identified Pro$_3^-$ to have a *cis* conformation of both peptide bonds, and AHA$^-$ to have a relatively low energy structure (~16 kJ/mol) with the first peptide bond in the *cis* configuration. This suggests that analogous to what has been suggested for protonated peptides, a *cis* configured peptide bond between the first and second residues is also a necessary prerequisite for diketopiperazine formation in anionic peptides.

Acknowledgments

The authors gratefully acknowledge Dr. D. Löwik from the Radboud University for synthesizing the Pro_2, Pro_6 and Pro_7 peptides. They gratefully acknowledge the FELIX staff, particularly Drs. G. Berden, A.F.G. van der Meer and B. Redlich for technical support. Financial support for this project was provided by NWO Chemical Sciences under VICI project nr. 724.011.002. The authors also thank NWO Physical Sciences (EW) and the SARA Supercomputer Center for providing the computational resources. This work is part of the research program of FOM, which is financially supported by NWO.

REFERENCES

(1) Aebersold, R.; Goodlett, D. R. *Mass spectrometry in proteomics* Chem. Rev. **2001**, *101*, 269.

(2) Steen, H.; Mann, M. *The ABC's (and XYZ's) of peptide sequencing* Nat. Rev. Mol. Cell Biol. **2004**, *5*, 699.

(3) Yates, J. R. *Mass spectrometry and the age of the proteome* J. Mass Spectrom. **1998**, *33*, 1.

(4) Ruedi, A.; Matthias, M. *Mass spectrometry-based proteomics* Nature **2003**, *422*, 198.

(5) Barlow, C. K.; O'Hair, R. A. J. *Gas-Phase Peptide Fragmentation: How Understanding the Fundamentals Provides a Springboard to Develping New Chemistry and Novel Proteomics Tools* J. Mass Spectrom. **2008**, *43*, 1301.

(6) Huang, Y.; Tseng, G. C.; Yuan, S.; Pasa-Tolic, L.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *A Data-Mining Scheme for Identifying Peptide Structural Motifs Responsible for Different MS/MS Fragmentation Intensity Patterns* J. Proteome Res. **2007**, *7*, 70.

(7) Li, W.; Ji, L.; Goya, J.; Tan, G.; Wysocki, V. H. *SQID: An Intensity-Incorporated Protein Identification Algorithm for Tandem Mass Spectrometry* J. Proteome Res. **2011**, *10*, 1593.

(8) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *Spectroscopic and theoretical evidence for oxazolone ring formation in collision-induced dissociation of peptides* J. Am. Chem. Soc. **2005**, *127*, 17154.

(9) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *Infrared spectroscopy and theoretical studies on gas-phase protonated Leu-enkephalin and its fragments: Direct experimental evidence for the mobile proton* J. Am. Chem. Soc. **2007**, *129*, 5887.

(10) Yoon, S. H.; Chamot-Rooke, J.; Perkins, B. R.; Hilderbrand, A. E.; Poutsma, J. C.; Wysocki, V. H. *IRMPD spectroscopy shows that AGG forms an oxazolone* b_2^+ *ion* J. Am. Chem. Soc. **2008**, *130*, 17644.

(11) Verkerk, U. H.; Siu, C. K.; Steill, J. D.; El Aribi, H.; Zhao, J.; Rodriguez, C. F.; Oomens, J.; Hopkinson, A. C.; Siu, K. W. M. a_2 *Ion Derived from Triglycine: An* N_1*-Protonated 4-Imidazolidinone* J. Phys. Chem. Lett. **2010**, *1*, 868.

(12) Bythell, B. J.; Molesworth, S.; Osburn, S.; Cooper, T.; Paizs, B.; Van Stipdonk, M. *Structure and Reactivity of* $a_{(n)}$ *and* $a_{(n)}(*)$ *Peptide Fragments Investigated Using Isotope Labeling, Tandem Mass Spectrometry, and Density Functional Theory Calculations* J. Am. Soc. Mass. Spectrom. **2008**, *19*, 1788.

(13) Perkins, B. R.; Chamot-Rooke, J.; Yoon, S. H.; Gucinski, A. C.; Somogyi, Á.; Wysocki, V. H. *Evidence of Diketopiperazine and Oxazolone Structures for HA* b_2^+ *Ion* J. Am. Chem. Soc. **2009**, *131*, 17528.

(14) Wang, D.; Gulyuz, K.; Stedwell, C. N.; Polfer, N. C. *Diagnostic NH and OH Vibrations for Oxazolone and Diketopiperazine Structures:* b_2 *from Protonated Triglycine* J. Am. Soc. Mass Spectrom. **2011**, *22*, 1197.

(15) Grzetic, J.; Oomens, J. *Spectroscopic identification of cyclic imide* b_2*-ions from peptides containing Gln and Asn residues* J. Am. Soc. Mass Spectrom. **2013**, *24*, 1228.

(16) Tirado, M.; Rutters, J.; Chen, X.; Yeung, A.; van Maarseveen, J.; Eyler, J. R.; Berden, G.; Oomens, J.; Polfer, N. C. *Disfavoring Macrocycle b Fragments by Constraining Torsional Freedom: The "Twisted" Case of QWFGLM* b_6 J. Am. Soc. Mass Spectrom. **2012**, *23*, 475.

(17) Chen, X.; Steill, J. D.; Oomens, J.; Polfer, N. C. *Oxazolone versus macrocycle structures for Leu-Enkephalin* b_2*-*b_4*: Insights from infrared multiple-photon dissociation spectroscopy and gas-phase hydrogen/deuterium exchange* J. Am. Soc. Mass Spectrom. **2010**, *21*, 1313.

(18) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *Mobile and localized protons: A framework for understanding peptide dissociation.* J. Mass Spectrom. **2000**, *35*, 1399.

(19) Paizs, B.; Suhai, S. *Combined quantum chemical and RRKM modeling of the main fragmentation pathways of protonated GGG. I. Cis-trans isomerization around protonated amide bonds* Rapid Commun. Mass Spectrom. **2001**, *15*, 2307.

(20) Paizs, B.; Suhai, S. *Combined quantum chemical and RRKM modeling of the main fragmentation pathways of protonated GGG. II. Formation of* $b_{(2)}$*,* $y_{(1)}$*, and* $y_{(2)}$ *ions* Rapid Commun. Mass Spectrom. **2002**, *16*, 375.

(21) Armentrout, P. B.; Heaton, A. L. *Thermodynamics and mechanisms of protonated diglycine decomposition: a computational study* J. Am. Soc. Mass Spectrom. **2012**, *23*, 621.

(22) Gucinski, A. C.; Chamot-Rooke, J.; Nicol, E.; Somogyi, Á.; Wysocki, V. H. *Structural Influences on Preferential Oxazolone versus Diketopiperazine b2+ Ion Formation for Histidine Analogue-Containing Peptides* J. Phys. Chem. A **2012**, *116*, 4296.

(23) Pal, D.; Chakrabarti, P. *Cis peptide bonds in proteins: residues involved, their conformations, interactions and locations* Journal of Molecular Biology **1999**, *294*, 271.

(24) Jabs, A.; Weiss, M. S.; Hilgenfeld, R. *Non-proline cis peptide bonds in proteins* Journal of Molecular Biology **1999**, *286*, 291.

(25) Grewal, R. N.; El Aribi, H.; Harrison, A. G.; Siu, K. W. M.; Hopkinson, A. C. *Fragmentation of Protonated Tripeptides: The Proline Effect Revisited* J. Phys. Chem. B **2004**, *108*, 4899.

(26) Bleiholder, C.; Suhai, S.; Harrison, A.; Paizs, B. *Towards Understanding the Tandem Mass Spectra of Protonated Oligopeptides. 2: The Proline Effect in Collision-Induced Dissociation of Protonated Ala-Ala-Xxx-Pro-Ala (Xxx = Ala, Ser, Leu, Val, Phe, and Trp)* J. Am. Soc. Mass Spectrom. **2011**, *22*, 1032.

(27) Breci, L. A.; Tabb, D. L.; Yates, J. R.; Wysocki, V. H. *Cleavage N-Terminal to Proline: Analysis of a Database of Peptide Tandem Mass Spectra* Anal. Chem. **2003**, *75*, 1963.

(28) Harrison, A. G.; Young, A. B. *Fragmentation reactions of deprotonated peptides containing proline. The proline effect* J. Mass Spectrom. **2005**, *40*, 1173.

(29) Andreazza, H. J.; Wang, T.; Bilusich, D.; Hoffmann, P.; Bowie, J. H. *Negative ion fragmentations of deprotonated peptides containing post-translational modifications: diphosphorylated systems containing Ser, Thr and Tyr. A characteristic phosphate/phosphate cyclisation. A joint experimental and theoretical study* Rapid Commun. Mass Spectrom. **2009**, *23*, 1825.

(30) Bowie, J. H.; Brinkworth, C. S.; Dua, S. *Collision-induced fragmentations of the* $(M\text{-}H)^-$ *parent anions of underivatized peptides: An aid to structure determination and some unusual negative ion cleavages* Mass Spectrom. Rev. **2002**, *21*, 87.

(31) Pu, D.; Clipston, N. L.; Cassady, C. J. *A comparison of positive and negative ion collision-induced dissociation for model heptapeptides with one basic residue* J. Mass Spectrom. **2010**, *45*, 297.

(32) Harrison, A. G.; Young, A. B. *Fragmentation reactions of deprotonated peptides containing aspartic acid* Int. J. Mass Spectrom. **2006**, *255-256*, 111.

(33) Oomens, J.; Steill, J. D. *The structure of deprotonated tri-alanine and its* a_3 *fragment anion by IR spectroscopy* J. Am. Soc. Mass Spectrom. **2010**, *21*, 698.

(34) Harrison, A.; Young, A. *Fragmentation of deprotonated N-benzoylpeptides: Formation of deprotonated oxazolones* J. Am. Soc. Mass. Spectrom. **2004**, *15*, 446.

(35) Harrison, A. G.; Siu, K. W. M.; El Aribi, H. *Amide bond cleavage in deprotonated tripeptides: a newly discovered pathway to "*b_2 *ions* Rapid Commun. Mass Spectrom. **2003**, *17*, 869.

(36) Grzetic, J.; Oomens, J. *Spectroscopic evidence for an oxazolone structure in anionic b-type peptide fragments* J. Am. Soc. Mass Spectrom. **2012**, *23*, 290.

(37) Harrison, A. G. *Sequence-specific fragmentation of deprotonated peptides containing H or alkyl side chains* J. Am. Soc. Mass Spectrom. **2001**, *12*, 1.

(38) Li, Z.; Yalcin, T.; Cassady, C. J. *C-terminal amino acid residue loss for deprotonated peptide ions containing glutamic acid, aspartic acid, or serine residues at the C-terminus* J. Mass Spectrom. **2006**, *41*, 939.

(39) Men, L.; Wang, Y. *Fragmentation of the deprotonated ions of peptides containing cysteine, cysteine sulfinic acid, cysteine sulfonic acid, aspartic acid, and glutamic acid* Rapid Commun. Mass Spectrom. **2006**, *20*, 777.

(40) Pu, D.; Cassady, C. J. *Negative ion dissociation of peptides containing hydroxyl side chains* Rapid Commun. Mass Spectrom. **2008**, *22*, 91.

(41) Harrison, A. *Effect of phenylalanine on the fragmentation of deprotonated peptides* J. Am. Soc. Mass. Spectrom. **2002**, *13*, 1242.

(42) Valle, J. J.; Eyler, J. R.; Oomens, J.; Moore, D. T.; van der Meer, A. F. G.; von Helden, G.; Meijer, G.; Hendrickson, C. L.; Marshall, A. G.; Blakney, G. T. *Free electron laser-Fourier transform ion cyclotron resonance mass spectrometry facility for obtaining infrared multiphoton dissociation spectra of gaseous ions* Review of Scientific Instruments **2005**, *76*.

(43) D.A. Case, T. A. D., T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Goetz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman *AMBER 2012* **2012**.

(44) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Wallingford CT, 2009.

(45) Oomens, J.; Steill, J. D.; Redlich, B. *Gas-phase IR spectroscopy of deprotonated amino acids* J. Am. Chem. Soc. **2009**, *131*, 4310.

(46) Bache, N.; Rand, K. D.; Roepstorff, P.; Ploug, M.; Jørgensen, T. J. D. *Hydrogen Atom Scrambling in Selectively Labeled Anionic Peptides Upon Collisional Activation by MALDI Tandem Time-of-Flight Mass Spectrometry* J. Am. Soc. Mass. Spectrom. **2008**, *19*, 1719.

Summary

The extraordinary functional diversity of proteins derives from the chemical uniqueness of the side chains of the amino acids, which directs the folding of the polypeptide chain in specific and unique architectures.

NMR spectroscopy and X-ray crystallography can determine the three-dimensional structural model of a protein. However, applications of these approaches are limited by protein size, conformational flexibility and aggregation predisposition. To provide complementary information about protein structure other approaches are required.

The discipline of genomics had an enormous impact on solving protein structures and functions. Development of DNA sequencing methods and techniques for extracting genes made it possible to define the amino acid sequence of a protein by determining the sequence of nucleotides in the corresponding coding gene.

If we now just observe a typical bacterium cell, it requires more than a few thousands of proteins in its life cycle. However, not all of these proteins are created at one moment and many are made under special conditions such as in specific new environments or if the cell is exposed to stress. The set of proteins existing in the cell, in a given space and time, represents its proteome. Proteomics is the study of the architecture, abundance, functionality and interactions of proteins and their assemblies which direct the activities of each living cell. Why proteomics? Complete information cannot be obtained from the study of genes alone; it is impossible to elucidate mechanisms of disease, aging, and effects of the environment or mechanisms of posttranslational protein modifications solely by studying the genome.

Mass spectrometry is one of the main tools in proteomics. Simply, in mass spectrometry experiments as applied to proteomics, peptide fragmentation is induced by, for example, collisions which results in cleavage of the peptide. This normally occurs primarily through one of the lowest energy pathways — that is, breaking of the peptide backbone bonds. From these fragments, ideally, one should be able to reconstruct the original amino acid sequence of a given peptide/protein, much like a puzzle.

In order to interpret peptide fragmentation spectra, peptide identification algorithms are used, offered by various software packages. Often, data is incompletely used; not all peptide mass fragments are defined, some may be miss-assigned and the information about the complex fragmentation chemistry of peptides is simplified in nearly all algorithms.

Knowledge of the fragmentation mechanisms and structures of fragment ions in the gas phase can improve our knowledge of chemistry in the absence of solvent. Additionally, such information can be used to show how different amino acid side chains affect the outcome of a competing dissociation reaction pathway. Recently, IR photo-dissociation spectroscopy of mass-selected peptide fragments, in combination with computational chemistry, has proven to be a very powerful tool to determine the molecular structures of these gas-phase reaction products.

The aim of experiments of the research described in this thesis was to advance the understanding of the fragmentation mechanisms and to establish new patterns within the set of these mechanisms for both protonated and deprotonated peptides. By recording the IR spectrum of a mass selected peptide fragment and assigning its structure by comparing the experimental spectrum with theoretical ones, we can suggest the fragmentation mechanism which might, or might not depend on the peptide conformation or substituents. The broader idea is, of course, to provide additional rules for interpreting peptide mass fragmentation spectra to be incorporated in existing algorithms that would increase the peptide sequence coverage and in general improve protein identification.

Chapter 1 gives the basics on protein, peptide and amino acid structure, how and why peptides fragment in mass spectrometry experiments and how such information can be used in identifying protein structures.

Chapter 2 contains a description and theoretical background of the experimental setup used for performing the studies in this thesis. Configuration of the FTICR mass spectrometer, the free electron laser FELIX and the IR-OPO are reported here.

Chapter 3 explains the fundamentals of IR/IRMPD spectroscopy, collision induced dissociation and briefly, theoretical approaches.

Protein identification using mass spectrometry is mainly based on analyzing protonated species. However, sequence coverage can readily be increased by analyzing the same peptide in both positive and negative ion mode and making use of the often complementary information. On modern MS instruments, switching from positive to negative ion mode is quick and easy. Still, information on the deprotonated peptides is poorly applied in practical proteomics applications, which is perhaps partly due to the fragmentation processes being less extensively studied. Thus, in Chapters 4, 5, 6 and 9 the focus is on the fragmentation chemistry of deprotonated peptides, showing that the "protonated trend" in proteomics can, and should be, modified.

Chapter 4 shows specifically how sequence coverage can be increased by analyzing the same peptide in both positive and negative ion modes. The c-type fragments generated from anionic peptides provide information that is not available from their positive-ion counterparts. Their identification can be used to determine the sequence from the N-terminus and thus improve sequence coverage. A set of c-type anionic peptide fragments is identified to be linear peptide fragments, deprotonated at various molecular sites depending on their gas phase acidities.

In **Chapter 5** experiments showing similarity in fragmentation of deprotonated peptides with their protonated analogs are presented. The b_2-type peptide fragments generated from deprotonated Ala_3 and AlaTyrAla are shown to possess oxazolone structure, well-known for b-type fragments of protonated peptides. The slightly acidic and bulky tyrosine amino acid residue does not affect formation of the oxazolone cyclic structure nor does it become the deprotonation site of the structure formed.

As in any other chemical reaction, gas phase fragmentation reactions also involve interaction of nucleophiles and electrophiles. Both can be the part of the peptide backbone, but they can also be constituents of some amino acids' side chains, such as in asparagine and glutamine. The question here is: can and how does the presence of a side chain nucleophile affect the fragmentation chemistry?

Chapter 6 concerns the fragmentation chemistry of deprotonated peptides containing asparagine. Results presented here are direct evidence of the side chain influence in fragmentation chemistry; the asparagine side chain nitrogen is the active nucleophile for the anionic cyclic succinimide b-type fragment formation.

Chapter 7 presents experiments analyzing the Gln and Asn side chain effect on the dissociation chemistry when the peptide is protonated. Similar to their anionic counterparts, b_2-type peptide fragments generated from these peptides possess a cyclic imide structure.

Due to the high dielectric constant of water and the major influence of this on electrostatic interactions, there is generally a significant difference between the conformation that a (ionized) peptide adopts in the gas phase and in aqueous solution. The strong interactions with water molecules will change the type and number of energetically advantageous structures that the molecule can assume. In the gas phase, and in particular in vacuum, the molecule is entirely isolated and is stabilized by intramolecular (electrostatic) interactions. The possible influence of the conformation of a peptide on its fragmentation behavior is examined in chapters 8 and 9 for peptides consisting of proline (Pro) containing peptides. Both positively and negatively charged peptides have been investigated.

In **Chapter 8** a series of protonated polyproline conformers having their first peptide bond in the *cis* conformation has been identified. Proline is unique among the common amino acids due to its ring structure, the secondary amine at the N-terminus and the rigidity it imparts on the peptide backbone when present. This molecular conformation of polyprolines is consistent with the unique fragmentation pattern upon collisional or multiple photon activation. Here, observation of the uncommon, but thermodynamically favored, diketopiperazine-type b_2^+ fragment ion is reported, which is in line with a *cis* configuration of the first amide linkage in the protonated parent ion.

Chapter 9 contains the results of an investigation on deprotonated polyproline peptides. Similar to its protonated analogue, the identified deprotonated tri-proline conformation also features its first peptide bond in cis configuration. This leads to the formation of diketopiperazine b_2-type peptide fragment. Further, the b_2-type fragment from deprotonated AlaHisAla is also a diketopiperazine. This is the first observation of the diketopiperazine pathway for anionic peptide fragmentation.

Clearly, infrared spectroscopy of gas phase molecular ions in combination with computational chemistry provides a wealth of information to elucidate molecular structures, which is in many cases not available from MS data alone. In conclusion, this thesis demonstrates that a deeper insight into molecular structures formed in ion fragmentation can help us understand mechanisms and kinetics of the dissociation processes in gas phase chemistry. This understanding, if extrapolated to higher domains, can be used to establish fragmentation rules and warrants improvements in protein identification in proteomics.

Samenvatting

De buitengewoon diverse functionaliteit van eiwitten valt te herleiden tot de unieke chemische eigenschappen van de 20 natuurlijk voorkomende aminozuurresiduen, die aanleiding geven tot zeer specifieke en unieke 3-dimensionale structuren van eiwitten. NMR spectroscopie en röntgendiffractie worden gebruikt om een 3-D model van de eiwitten te verkrijgen, hoewel deze methoden niet altijd toepasbaar zijn; de grootte van eiwitten, hun conformationele flexibiliteit en hun wateroplosbaarheid kunnen de toepassing van deze methoden beperken. Verschillende andere methoden worden daarom ingezet om complementaire informatie te verschaffen.

Het bepalen van de structuur en functie van eiwitten heeft een vlucht genomen met het in kaart brengen van de genetische informatie (genomics). De ontwikkeling van methoden om DNA te sequencen heeft het mogelijk gemaakt de volgorde van aminozuren in een eiwit, de zogenaamde primaire structuur van het eiwit, te achterhalen via de volgorde van de nucleotiden die voor het DNA coderen.

Zelfs een eencellig organisme zo als een bacterie gebruikt al duizenden eiwitten in haar levenscyclus. Niet al deze eiwitten komen tegelijkertijd tot expressie en sommigen eiwitten zelfs alleen onder bepaalde omstandigheden, beïnvloed door bijvoorbeeld omgevingsfactoren of stress. De studie van de architectuur, relatieve concentraties, functionaliteit, en interacties van de gehele verzameling eiwitten van een organisme, het proteoom, wordt aangeduid met de term *proteomics*. Waarom is proteomics belangrijk? De complete werking van een cel kan niet worden afgeleid uit kennis van het genoom alleen; belangrijke mechanismen zoals ziektes, veroudering, en omgevingsinvloeden hangen samen met post-translationele modificaties van de eiwitten en kunnen dus niet uit het genoom worden afgeleid.

Massaspectrometrie is een van de belangrijkste analytische methodes in proteomics. In het kort, uit het moleculair gewicht van een eiwit (MS) en van zijn moleculaire fragmenten (MS/MS) – waarbij fragmentatie bijvoorbeeld kan worden geïnduceerd door botsingen in de massaspectrometer – kan men de volgorde (Engels: sequence) van aminozuurresiduen in het eiwit reconstrueren. Fragmentatiereacties verlopen in het algemeen langs de laagst energetische reactiepaden; eiwitten fragmenteren daarom voornamelijk door het breken van de peptidebinding, de link tussen twee residuen. Door het oplossen van de puzzel die onstaat na het fragmenteren van het eiwit, kan men dus de primaire structuur van het eiwit bepalen.

In het algemeen worden computeralgoritmes gebruikt om de volgorde van residuen te reconstrueren uit een gemeten MS/MS spectrum. Deze algoritmes gebruiken echter slechts een klein deel van de informatie die verborgen ligt in het massaspectrum; van veel pieken is de identiteit niet bepaald of mogelijk verkeerd bepaald. Bovendien wordt het complexe mechanisme van de dissociatiereactie niet of zeer versimpeld in beschouwing genomen.

Informatie over de moleculaire structuren van de gevormde fragmenten en van de kinetiek van de dissociatiereacties kan ons begrip van de reactiemechanismen in de gasfase sterk verbeteren. Zo kan deze informatie bijvoorbeeld gebruikt worden om te bezien welke invloed de verschillende residuen hebben op het reactiepad van de dissociatie. Infrarood fotodissociatie-spectroscopie van massa-geselecteerde fragmentatieproducten is een krachtig instrument om de precieze moleculaire structuur van die fragmenten te bepalen.

Het doel van het onderzoek dat beschreven is in dit proefschrift is onze kennis van dissociatiemechanismen te vergroten en eventueel nieuwe dissociatieroutes te ontdekken voor zowel geprotoneerde als gedeprotoneerde peptides. Door IR spectra op te nemen van specifieke massa-geselecteerde peptidefragmenten en vervolgens hun structuur te bepalen door vergelijking van experimentele en theoretische IR spectra, kunnen we reactiemechanismen suggereren die mogelijk afhankelijk zijn van de conformatie of constituenten van het oorspronkelijke peptide. Het dieper liggende doel is om te komen tot een verbeterde set regels om MS/MS spectra van peptides te interpreteren. Die regels kunnen in toekomstige algoritmes worden toegepast om zo uiteindelijk eiwitidentificatie te verbeteren.

Hoofdstuk 1 behandelt enkele basisaspecten van de studie naar de structuur van aminozuren, peptides en eiwitten, hoe en waarom peptides dissociëren in een massaspectrometer, en hoe deze informatie gebruikt wordt om eiwitten te identificeren in proteomics.

Hoofdstuk 2 geeft een beschrijving en theoretische achtergrond van de experimentele instrumenten en methodes die gebruikt zijn voor de studies in dit proefschrift. De opzet van de Fourier Transform ion cyclotron resonance (FTICR) massaspectrometer, de vrije-electronenlaser FELIX en het infrarood lasersysteem gebaseerd op optische parametrische oscillatie (OPO) worden hier toegelicht.

Hoofdstuk 3 behandelt enkele fundamentele aspecten van IR spectroscopie en in het bijzonder IR multifoton-dissociatie (IRMPD) spectroscopie. Voorts worden botsingsgeïnduceerde dissociatie (CID) in een massaspectrometer en de quantum-chemische methodes die in dit proefschrift zijn gebruikt kort toegelicht.

Identificatie van eiwitten met behulp van massaspectrometrie wordt vooral uitgevoerd door geprotoneerde (positief geladen) peptides te analyseren. Het aantal unieke, sequence-bepalende, fragmenten in een MS/MS spectrum, de zogenaamde *sequence coverage*, kan vaak eenvoudig vergroot worden door dezelfde peptide ook te analyseren als negatief ion en gebruik te maken van complementaire informatie in de twee MS/MS spectra. Moderne massaspectrometers kunnen gemakkelijk en snel schakelen tussen positieve en negatieve ionisatie. Niettemin wordt informatie van gedeprotoneerde (negatief geladen) peptides slechts zeer zelden gebruikt in typische proteomics studies. Dit is wellicht ook te wijten aan de veel gebrekkigere kennis van de dissociatiereacties van gedeprotoneerde peptides. In hoofdstukken 4, 5, 6 en 9 van dit proefschrift bestuderen daarom de dissociatiereacties van gedeprotoneerde peptides en laten zien dat deze methodes inderdaad complementaire informatie kunnen verschaffen.

Hoofdstuk 4 laat heel spcifiek zien hoe MS/MS experimenten aan gedeprotoneerde peptides aanvullende informatie verschaffen ten opzichte van de informatie uit MS/MS spectra van het positieve ion. De c-type fragmenten die alleen geproduceerd worden door negatief geladen peptides bevatten informatie die niet beschikbaar is uit de positief geladen peptides. Identificatie van deze c-type fragmenten kan gebruikt worden om het peptide te sequencen vanaf de N-terminus en kan zodoende de sequence coverage vergroten. Onze experimenten tonen aan dat de c-type fragmenten een lineaire moleculaire structuur bezitten en dat de positie van deprotonering afhangt van de aciditeit (zuurgraad) in de gasfase van de residuen in het peptide.

In **hoofdstuk 5** worden overeenkomsten in de fragmentatie van gedeprotoneerde en geprotoneerde peptides onderzocht. We tonen aan dat de b_2-type fragmenten van de gedeprotoneerde peptides Ala_3 en AlaTyrAla een oxazolone-structuur bezitten, die bekend is voor b-type fragmenten van geprotoneerde peptides. De verhoogde zuurgraad van het tyrosine residue beïnvloedt de vorming van de oxazolone-ring niet en heeft ook geen invloed op de uiteindelijke positie van deprotonering.

Zoals in vrijwel alle chemische reacties spelen interacties tussen nucleofielen en electrofielen ook in dissociatiereacties in de gasfase een belangrijke rol. Beiden kunnen onderdeel zijn van de ruggengraat van de peptide, maar zij kunnen zich ook in de residuen van de aminozuren bevinden. De amide-groepen in de residuen van glutamine (Gln) en asparagine (Asn) vormen bijvoorbeeld sterk nucleofiele groepen. De vraag rijst dan of en hoe de aanwezigheid van zulke groepen in de residuen de dissociatiereacties beïnvloeden.

Hoofdstuk 6 beschrijftde fragmentatiereacties van gedeprotoneerde peptides die een Asn residu bevatten. De experimenten tonen onomstotelijk aan dat het asparagine residu direct deelneemt aan de omleggingsreactie die tot dissociatie leidt; het stikstofatoom uit de amidegroep van het residu blijkt de actieve nucleofiel te zijn, waardoor het anionische b-fragment een cyclische succinimide structuur aanneemt.

Hoofdstuk 7 presenteert een studie naar de effecten van Gln en Asn residuen op de dissociatiereacties van geprotoneerde (positief geladen) peptides. Analoog aan de situatie in anionische peptides vinden we hier dat de b_2-type fragmenten een cyclische imide structuur bezitten, in tegenstellling tot de welbekende oxazolone structuur.

Door de hoge diëlectrische constante van water en de grote invloed hiervan op electrostatische interacties, is er over het algemeen een groot verschil tussen de conformatie die een (geïoniseerd) peptide aanneemt in de gasfase en in een waterige oplossing. De sterke interactie met watermoleculen verandert het type en aantal energetisch voordelige structuren die het molecuul kan aannemen. In de gasfase, in het bijzonder in vacuüm, is het molecuul volledig geïsoleerd en wordt gestabiliseerd door intramoleculaire (electrostatische) interacties. De mogelijke invloed van de conformatie van de peptide op haar fragmentatiegedrag wordt in hoofdstukken 8 en 9 onderzocht voor peptides die bestaan uit proline (Pro) residuen. Zowel positief als negatief geladen peptides zijn onderzocht.

In **hoofdstuk 8** zijn de conformaties van een serie polyproline peptides (Pro_n) geïdentificeerd waaruit blijkt dat in deze polypeptides de eerste peptidebinding (gerekend vanaf de N-terminus) gewoonlijk in de *cis*-conformatie voorkomt, die voor de meeste peptides erg ongebruikelijk is. Van alle aminozuren is Pro bijzonder vanwege zijn ringvormige structuur en zijn secundaire (in plaats van primaire) amine; peptides met een Pro residu vertonen een ongebruikelijke rigiditeit ter plaatse van het Pro residue. Peptides die alleen uit Pro-residuen bestaan, nemen bijzondere secunadaire structuren aan die ook leiden tot bijzondere fragmentatiepatronen in MS/MS spectra. In dit hoofdstuk tonen we door middel van IR ionspectroscopie aan dat de b_2-type fragmenten van deze polyprolines een diketopiperazine structuur hebben. Hoewel deze structuur thermodynamisch stabieler is, komt zij weinig voor omdat de oxazolone-structuur kinetisch voordeliger gevormd kan worden. Echter, de *cis*-conformatie van de peptidebinding geeft een natuurlijke verklaring voor de vorming van de diketopiperazine-structuur in dit bijzondere geval.

Hoofdstuk 9 is gewijd aan gedeprotoneerde polyproline peptides. Analoog aan de bevindingen voor de geprotoneerde systemen, vinden we voor gedeprotoneerd Pro_3 een conformatie waarin de eerste peptidebinding in een *cis*-structuur voorkomt. Ook hier leidt dat na dissociatie tot een b_2-type fragment met een diketopiperazine structuur. Daarnaast zijn gedeprotoneerde peptides met een histidine residue (His) onderzocht, waarvoor bekend is dat de geprotoneerde versie een diketopiperazine structuur vertoont. Ook het anionische b_2-type fragment van gedeprotoneerd AlaHisAla bezit deze bijzonder structuur. Deze experimenten tonen voor het eerst aan dat de diketopiperazine structuur ook in negatief geladen peptides gevormd kan worden.

Samenvattend, infrarood spectroscopie aan gasvormige moleculaire ionen gecombineerd met quantum-chemische berekeningen verschaft een zee aan informatie over de moleculaire structuur van de ionen, die in de meeste gevallen niet of niet eenduidig afleidbaar is uit massa-spectrometrische metingen alleen. Dit proefschrift laat in het bijzonder zien dat een beter begrip van de moleculaire structuren die gevormd worden tijdens de dissociatie van geïoniseerde peptides, ons kan helpen een beter inzicht te krijgen in de mechanismen en kinetiek van de dissociatie-chemie. Dit verbeterde inzicht kan uiteindelijk gebruikt worden om MS-gebaseerde identificatie van eiwitten te verbeteren.

Acknowledgments

I remember it was June 25th, 2010, a sunny day in Ijsselstein during the football championship when I met my supervisor Jos Oomens for the PhD interview. I remember that the day was so lively, with tulips, and everything in orange.

Jos, thank you so much for the trust you gave me and that you believed I was capable of performing this task and that I deserve the title that I will (hopefully) start carrying in a couple of weeks from now. I appreciate your exceptional scientific supervision, personal guidance and your friendly attitude. I am honored that I had opportunity to work and learn from you!

In the Rijnhuizen castle it was easy to feel like a king. There I met Giel Berden, a person who I could ask anything and get an answer. Thank you Giel for your patience, for sharing your knowledge, for always having time for me, for helping me whenever I had problems, for Kinderdijk and for being a true friend!

Further on, I hold great appreciation for ex and present members of the FELIX team, Lex van der Meer, Rene van Buuren, Michel Riet, Guus Tielemans, Theo Neuij, Jan Pluijgers, Wybe Roodhuyzen , Jules van Leusden and especially for Britta Redlich who, as a role model, was always impressing me with her strength, persistence, organizational skills and kindness! Thank you all for being so helpful in providing the perfect light for my experiments.

I am very grateful to Anouk Rijs for being a kind room and officemate. Dear Anouk, I hope that life will treat you pink!

I also highly appreciate the influence of my co-supervisor Wybren Jan Buma. Thank you for the encouragements and for contributing to the diversity of my research.

My dear comrades in work as well as in beer drinking Sander Jaeqx, Denis Kiawi, Vivike Lapoutre, Daniël Bakker, Joost Bakker, Lisanne Kempkes, Gydo van Zundert, Jordy Bouwman, Nikolas Stavrias and Hector Alvaro Galue, thank you!

However, the most memorable time in the lab I had with Mitra Almasian. Mitra, thank you for being a friend and for sharing similar interests, taste in music and a great sense of humor.

Being in RIjnhuizen was adorn by gezellig atmosphere. I always appreciated the charming and warm welcome of Loes van de Ven, An Uwland and Jose Kragten at the work entrance. Thank you girls so much! They are not the only ones that will be missed. People that I always enjoyed to see and spend time with are Frits Hekkenberg, dear Caroline, Miranda van den Berg and unforgettable Noortje Khan. Hartelijke bedankt to all!

When we moved to Radboud University in 2012, I thought how sad it was. Not just that moving was painful, difficult and relatively unappreciated to a PhD project at the halfway point of research, but I also felt the emotional gap - I was missing some of the Rijnhuizen people! That gap was filled with a newly met, lovely and kind person - Miriam Heijmerink. Thank you Miriam for smiles, warmth and for the chocolate! It all meant a lot to me!

Here, I would also like to thank to Wim van de Zande for sharing his knowledge and enthusiasm for science.

During my PhD I was lucky enough to end up communicating and sometimes collaborating with international visitors working at the FELIX facility. That broadened my horizons and enriched me, both professionally and personally. Therefore, a big thanks to Mike van Stipdonk, Udo Verkerk, Victor Ryzhov, Peter Armentrout, Rob Dunbar and Jeffrey Steill.

Hereby I would like specially to thank Nicolas Polfer and Erin Kristen for making my two months working in Florida easy, fun and unforgettable. Nick, thank you for being a friend as much as a supervisor during the time I spent in your lab!

Furthermore, I am very grateful to Jonh Eyler! Not just that he was working on building up the FTICR instrument that I was using in my experiments but he also lent his bike to me while I was in Florida. The possibility to cycle there was a true salvation.

Although writing a thesis is generally considered to be an individual task, comments, tips and corrections were highly appreciated and without them I think that this task would be almost impossible. Thus, thanks to the thesis reading committee members Britta Redlich, Dave Parker and Nicolas Polfer, and also for accepting the responsibility, for the spent time and of course for the approval. For the same reasons, thanks to my supervisor Jos Oomens, and my friends Giel Berden and Jonathan Martens.

There are also people at Belgrade University that I would like to say thanks to - Ilija Brceski and Jasminka Korolija. Thank you for support, understanding and for being an example of preserved intellect in a less than conscious society.

Now thanks to my great family and friends for being supportive, for all the fun, the sunny vacations and for the love! And thank you is just not enough! Mira Borozan ex Grzetic, Mato Grzetic, striko Grgo Grzetic, Nemanja Pemos Karan, David i Lena Pasinovic, Tomovic kum Slobo, (ah ta Ljuta), Dragana Taska Lukic, Jovana i Davor Dukic, Ana Milicevic, Marina, Vlada i Milja Dmitrovic, Bokica Djavolak Milovanovic, Jovica Plavsic i Svetlana, Nikola Vunjerovic, Aleksandra Miljevic, Milica Jez, Anja Potkonjak, Sladjana Vitezovic, Radmila Shapica Stojanovic, Relja Mirkovic, Nikola Zotovic, Milica Grujicic, Ivan Ilic, Dragutin Radicevic, Radojica Nikolic, Danka i Milos Labovic, Konstantin Ilijevic.

Amusingly, similar to Jesus I also have two sacred, graceful Marias - the small one Marija Mara Cvetkovic and the big one Marija kuma Grka Pasinovic, and yes, they always make me feel blessed.

Divine, perfect, the best and beautiful, my sister Dragica Lucija Toc Grzetic. Thank you dear sister, greatest of all sisters, without whom I`d never survive. Thank you for all the love and smiles you gave me, and more than everything, thank you for Danilo! (our hearts are irrevocably combined | across the sky until the end of time)

Dear, adorable, curly Jon Jonko Martens, well what can I say… thank you first of all for adding the “the” and “a” in this thesis, for support, protection, understanding, fun and above all - love. You came like a cherry on top of the cake of this four years. Thank you for being wonderful and awesome and making dreams come true. Happiness is only real when shared.

Mom and dad. Thank you for blessing me with this miraculous life and unconditional love, for continuous support and the possibilities to learn, work, laugh and love! Thank you for showing me that true strength lies in true love. Thank you for my sister! Thank you for the sea! Thank you for teaching me how to make the right choices! Thank you for each fairytale and space story! You made it all possible so that my life has become one of those tales! I will forever be your crvic I bubica! Volim vas beskrajno!

Josipa Grzetic
Nijmegen, 2014

Curriculum Vitae

Josipa Grzetic
born on 20th January 1980
Belgrade, Yugoslavia

1995-1999 **Gymnasium "IX gimnazija-Mihajlo Petrovic Alas", Belgrade, Republic of Serbia**
Graduate thesis "Nuclear reactions"

1999-2009 **University of Belgrade, Republic of Serbia**
Faculty of Chemistry, Department of Research and Development
M.Sc. in Chemistry
Graduate thesis "Synthesis and characterization of novel usnic acid derivative and its complex with Cu (II)"

2010-2014 **FOM Institute DIFFER/Radboud University Nijmegen, PhD studies**
Chemistry of peptide fragmentation - Molecular structures by ion spectroscopy

Supervisors: Prof. Jos Oomens (Radboud University Nijmegen)
Prof. Wybren Jan Buma (University of Amsterdam)

Publication List

1. **J. Grzetic**, J. Oomens, *Spectroscopic Evidence for an oxazolone Structure in Anionic b-Type Peptide fragments*, J. Am. Soc. Mass Spectrom. 23 (2012) 290-300, doi: 10.1007/s13361-011-0297-3

2. **J.Grzetic**, J. Oomens, *Structure of anionic c-type peptide fragments elucidated by IRMPD spectroscopy*, Int. J. Mass spectrom. 316-318 (2012) 216-226, http://dx.doi.org/10.1016/j.ijms.2012.02.012

3. **J. Grzetic**, J. Oomens, *Effect of the Asn side chain on the dissociation of deprotonated peptides elucidated by IRMPD spectroscopy*, Int. J. Mass Spectrom. 345-355 (2013) 70-77, doi: 10.1016/j.ijms.2013.05.005

4. **J. Grzetic**, J. Oomens, *Spectroscopic Identification of Cyclic Imide b_2-Ions from Peptides Containing Gln and Asn Residues,* J. Am. Soc. Mass Spectrom. 8 (2013) 1228-41, doi:10.1007/s13361-013-0661-6

5. P.Y. Iris Shek, J. Kai-Chi Lau, J. Zhao, **J. Grzetic**, U. H. Verkerk, J. Oomens, A. C. Hopkinson, K.W. M. Siu, *Fragmentations of protonated cyclic-glycylglycine and cyclic-alanylalanine*, Int. J. Mass Spectrom. 316–318 (2012) 199– 205, http://dx.doi.org/10.1016/j.ijms.2012.02.011

6. M. Almasian, **J. Grzetic**, G. Berden, B. Bakker, W. J. Buma, J. Oomens, *Gas-phase infrared spectrum of the anionic GFP-chromophore*, Int. J. Mass Spectrom. 330-332 (2012) 118-123, http://dx.doi.org/10.1016/j.ijms.2012.08.017

7. M. Almasian, **J. Grzetic**, J. van Maurik, J. D. Steill, G. Berden, S. Ingemann, W. J. Buma, J. Oomens, *Non-Equilibrium Isomer Distribution of the Gas-Phase Photoactive Yellow Protein Chromophore,* J. Phys. Chem. Lett. 3 (2012) 2259—2263, doi: 10.1021/jz300780t

8. A. Piatkivskyi, S. Osburn, K. Jaderberg, **J. Grzetic**, J. D. Steill, J. Oomens, J. Zhao, J. Kai-Chi Lau, U. H. Verkerk, A. C. Hopkinson, K. W. M. Siu, V. Ryzhov, *Structure and Reactivity of the Distonic and Aromatic Radical Cations of Tryptophan*, J. Am. Soc. Mass Spectrom. 24 (2013), 513-523, doi: 10.1007/s13361-013-0594-0

9. J. Zhao J, J.K. Lau, **J. Grzetic**, U. H. Verkerk, Oomens J, K. W. Siu, A. C. Hopkinson, *Structures of $a_{(n)}$* ions derived from protonated pentaglycine and pentaalanine: results from IRMPD spectroscopy and DFT calculations,* J. Am. Soc. Mass Spectrom. 24 (2013), 1957-1968, doi:10.1007/s13361-013-0728-4

10. J. K. Martens, **J. Grzetic**, G. Berden, J. Oomens, *Gas-phase conformations of small polyprolines and their fragment ions by IRMPD spectroscopy,* Int. J. Mass Spectrom. DOI: 10.1016/j.ijms.2014.07.027 (2014)

11. **J. Grzetic,** J. K. Martens, J. Oomens, *Observation of anionic diketopiperazine b_2 fragment ions: an IRMPD study,* (2014), manuscript in preparation